The Dissertation Committee for Todd Brandon Krause
certifies that this is the approved version of the following dissertation:

# Darwinian Evolution: the Mutation of a Weakly Relativistic Lagrangian

Committee:

Philip J. Morrison, Supervisor

Richard Hazeltine

Wendell Horton, Jr.

Richard Fitzpatrick

Irene Gamba

# Darwinian Evolution:
# the Mutation of a Weakly Relativistic Lagrangian

**by**

**Todd Brandon Krause, B.A.**

**DISSERTATION**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**DOCTOR OF PHILOSOPHY**

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2004

patri matrique

causa non cogendi ut consequar

gradum philosophiae doctoris

# Acknowledgments

Many people have assisted in the completion of this thesis, some willing, some unwilling. I wish to extend my thanks to those of both categories, though mentioning each by name would of course lead to unforseen omissions. Let this serve as a collective thank you, a sincere and heartfelt admission of my debt of gratitude to those who have struggled along with me to produce this work.

Some have made such outstanding contributions to this thesis that I cannot fail to mention them by name. First I must of course thank my advisor Phil Morrison, for whose patience and direction through these long years few words could convey my gratitude. Thanks also to the rest of the Morrison research group, particularly Alex Wurm for proofreading an earlier version of this thesis, and to Evstati Evstatiev for many useful discussions.

To Amit Apte go the thanks of a true and deep friendship, one which has benefitted me more than him, as he has spent countless hours explaining the intricacies of Physics to a mind slow to grasp them. *bahut bahut shukriā, kyõki tumhẽ mālūm hai ki is shodh-prbhand kā jo bhi acchā hai, vah tum se āta hai. mere fiziks ke guru haĩ āp, aur sab se acche mitra.* Thanks also to Ruth and Neal, for allowing me to steal Amit at inconvenient times.

Naturally my parents deserve a great share of the praise for completion of this thesis, for they were the ones who started me on my begrudging academic path, and they have supported me in so many ways along the journey.

Words cannot describe the thanks I feel towards Gladys Helena, and I hope that she will understand if here I say less to convey more.

I would be remiss if I failed to thank those who have supported my pursuits outside of Physics. Thanks to Nicholas Matlis, for being ‘da Man’ in so many ways; thanks to the whole ITF Taekwon-do Club, especially Mr. Travis Young and Asha Nadipuram. Thank you to Drs. Winfred Lehmann and Jonathan Slocum of the Linguistics Research Center for giving me a chance to hone my linguistic skills, albeit at a pace much slower than anyone would have hoped. Thanks to Mark Southern, for giving me a foundation in Historical Linguistics and the constant encouragement to pursue my little ideas; thanks to Doug Simms, Mary Blockley, and Ernie Kaulbach for coddling my forays into various languages. Thanks to Eric and Cheryl Kaufman for giving me an outlet for my grammatical diatribes.

Thanks to Matthew Deady, my advisor at Bard College, but more importantly my dear friend, who started me on the path to a Physics degree.

Forgive the omission of others who have helped directly and indirectly. It is certainly not for lack of recognition of their input — rather I am plagued by the fact that I am indebted to too many people, whose kind hearts and generosity should not go unmentioned, though to give them due credit would occupy yet another volume.

**Darwinian Evolution:**
**the Mutation of a Weakly Relativistic**
**Lagrangian**

Publication No. ____________________

Todd Brandon Krause, Ph.D.
The University of Texas at Austin, 2004

Supervisor: Philip J. Morrison

The work studies Darwin's order-$(v/c)^2$ approximation to the relativistic interaction of classical charged particles. The first part presents an introduction to the methods of symplectic reduction in the Newtonian two-body problem and then applies these to a two-body Darwin interaction. The momentum-dependent interaction of the Darwin system plays an important role in the ability to reduce to a system of one degree of freedom. Circular orbits are sought, and it is shown that two of the three possible orbits are prohibited by velocity conditions.

The second part of the work derives a self-consistent Darwin particle theory from a Lagrangian for electromagnetic fields coupled to particles. The resulting particle Lagrangian agrees with previous results. A similar procedure is followed to obtain a Low-Darwin system, coupling the self-consistent Darwin theory to the Vlasov equation.

# Table of Contents

# Chapter 1

# Introduction

The following chapters contain the highlights of an unfolding story, one which contains too many equations to turn a profit at the bookstands, but remains nonetheless filled with drama and intrigue, personal failures and successes, all thinly veiled behind the banner of scientific investigation. In many places the reader will find more mathematical detail than would strictly be necessary in a polished work of scientific research. This is deliberate, done with the hope of *instructing* the reader in the topic at hand, rather than merely displaying a polished work devoid of personal content. The author firmly believes that teachers are a necessary antecedant to researchers, that the former can continue without the latter, but not the latter without the former. This dissertation therefore hopes to teach a thing or two about the Darwin system of interacting charges, without too blatantly foisting itself upon the reader as the product of original research. If the reader gains from this work a deeper understanding of some traditional textbook material, then the author will consider the entire work a success.

## 1.1 Background

The story of the Darwin Lagrangian begins with C.G. Darwin's original paper "The Dynamical Motions of Charged Particles", published in 1920 [9]. Darwin's stated purpose was "to reduce the problem of the motion of any number of charged particles, moving at high velocities in an electric and magnetic field, to a Lagrangian form, so that all the well-known theorems of general dynamics may be made applicable." In practice this involved an order-$(v/c)^2$ approximation to the fully relativistic motion of charged particles. By working to this order, one eliminates the complexities of dealing with retarded quantities, and thus calculation becomes more practicable.

Darwin's basic method was essentially as follows. First, given the Lorentz force law for the motion of a particle in a given electromagnetic field, he recast the fields in terms of potentials and constructed from there a Lagrangian for the system. He then supposed the given fields to be produced by another charged particle. The potentials could thus be determined from the motion of this second particle. In particular, Darwin expressed the fields as the Liénard-Weichert potentials due to the second charge, which are given explicitly in terms of the coordinates and velocity of the particle. The Liénard-Weichert potentials, however, are written in terms of retarded times, and so Darwin expanded these to second order in $v/c$ and inserted the resulting expressions into the Lagrangian. A sum over particles completed the process for an $N$-body system.

Darwin then proceeded to apply the results to a problem of extreme interest at the time, the Bohr-Sommerfeld atom. The basic impetus from Darwin's point of view was that the methods then being applied to the atom were those of the Hamilton-Jacobi equation. This, being in terms of momenta, was deemed less direct for matters of experiment, and Darwin preferred to express the system in terms of coordinates and velocities, the more easily measurable quantities.

Since that time, the Darwin system has remained for the most part under the purview of researchers in quantum theory. Any search for the various applications of Darwin's approximation must inevitably be prepared to wade through masses of quantum literature before hitting upon any classical applications. Nevertheless there still are many classical studies applying Darwin's research, and the present thesis endeavors to contribute to this body of work.

Though originally formulated as a classical system, very little research has been done on the Darwin Lagrangian from a purely classical mechanical or dynamical systems point of view. Some, such as Dettwiller [12], have focused on the conserved quantities in the system. Presumably the paucity of other research in this vein comes from a general feeling that the system is 'too simple' in dynamical terms. But it is exactly its simplicity which might make it suitable for elucidating some of the finer points of modern dynamical systems theory. This is the motivation behind the application of symplectic reduction to the system presented in this thesis. The minor differences between the Darwin and Newtonian two-body problems allow one to see in bolder relief the limits of conclusions drawn from modern mathematical techniques.

By far the majority of literature using the Darwin system in a classical setting employs it as a simplifying approximation, in essence as it was originally intended. Several of these investigations are specifically applied to plasma physics. Appel and Alastuey [2, 3] have investigated the equilibrium properties of low-density, one-component plasmas governed by the Darwin interaction. Kaufman [31] applied the Darwin approximation to study the magnetic susceptibility of an imperfect gas. Kaufman and Rostler [32] in 1971 were among the first to propose the Darwin model as a self-standing basis of plasma simulation. Several have taken up the idea, and the Darwin model has found its greatest applications in Particle-in-Cell plasma simulation codes. Hewett [26] has given an overview of applications of the Darwin model to the simulation of low-frequency plasma phenomena: in Mirrortron simulations, simulating the fields of an ion-bunch accelerated by repulsion from a trailing tail of ions; in simulating instabilities arising in a plasma column with stationary ions and counter-streaming electron components. Pritchett [49] points out that, although the Darwin approximation works well in two-dimensional simulations, it suffers from difficulties in three dimensions with nonperiodic boundary conditions. The problem of boundary conditions has been taken up in some detail by Weitzner and Lawson [55].

## 1.2 Structure of the Thesis

The present work falls mainly into two parts, the first of a more geometrical and dynamical character, the second more typically physical. The first part comprises Chapter 2 and Chapter 3. The first of these chapters is an introduction to (or review of) the basics of symplectic reduction as they apply to the two-body problem of standard Newtonian mechanics. This is the archetypical example of the methods of reduction, so eloquently presented by Smale himself [52]. The treatment here is far more basic than Smale's, and is intended to be a makeshift bridge between the methods of physics and of symplectic topology. The following chapter then introduces the main object of study, the Darwin system. This is an order-$(v/c)^2$ approximation to the fully relativistic interaction of charged bodies, one which replaces the retarded-time interaction of relativity with an action-at-a-distance formulation amenable to the standard methods of Newtonian mechanics. The methods of symplectic reduction studied in reference to the Kepler problem are then co-opted and applied to a two-body system subject to the Darwin interaction. In the process, interesting questions arise concerning the procedure of symplectic reduction. From there focus is drawn to a study of the fixed points of the equations of motion and the conditions for circular orbits.

The second part of the work comprises Chapter 4, Chapter 5, and Chapter 6. These contain calculations relating to a different approach to the derivation of Darwin theory, in which Maxwell's equations are included in the action from the beginning, instead of being imposed as secondary constraints. Chapter 4 collects some preliminary notions relating to the Darwin system, so that the detail of subsequent explicit calculations will not obscure too much the forest for the trees. Chapter 5 applies this procedure to the $N$-body problem.

Chapter 6 takes the same action, with particles and fields, as its starting point, but proceeds to develop a theory coupled to the Vlasov equation using a formulation originally given by Low.

Several appendices add considerably to the heft of the book. Certainly a more streamlined dissertation could have done away with the material altogether. But several of the calculations were instructive enough to warrant inclusion if the author could find the time to type them up, and some were tedious enough that their inclusion is almost necessary to allow others to reproduce the results. Their inclusion in the main body of the text, however, would have distracted too much from the flow of the arguments, and so they have been relegated to appendices.

It is the author's sincere hope that what is old in this dissertation will nevertheless be found instructive, and what is new will be found by the reader nearly as exciting as it is to the author.

# Chapter 2

# A Primer on Symplectic Reduction: The Kepler Problem

This chapter discusses the well-known Kepler problem. The familiarity of the problem makes its reformulation in terms of the techniques of symplectic reduction less shocking. The physics at every stage should be transparent, so that one may more easily become accustomed to the mathematical structures being introduced.

## 2.1 The Two-Body Problem

First let us discuss the usual physics treatment of the problem, so that we may agree on notation. Consider two masses, $m_1$ and $m_2$. Nowhere in the following discussion will we use the positivity of the masses, so that the results will apply equally well to Coulomb interaction. The masses have phase coordinates $(\mathbf{r}_1, \mathbf{p}_1)$ and $(\mathbf{r}_2, \mathbf{p}_2)$, respectively. They interact through the central force

$$\mathbf{F}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{Gm_1m_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3}(\mathbf{r}_1 - \mathbf{r}_2) \ . \tag{2.1}$$

The system is described by the twelve coordinates $(\mathbf{r}_1, \mathbf{p}_1; \mathbf{r}_2, \mathbf{p}_2)$.

### 2.1.1 Conservation of Momentum

The principle of *conservation of momentum* allows us to choose a frame of reference in which the total momentum of the system is zero. This is a preferential reference frame, and we may now use this frame to give an "absolute" meaning to momentum in any other frame. If we pick any other frame moving at velocity $\mathbf{V}$ relative to the zero-momentum frame, then we may write the total momentum of the system in that frame as

$$\mathbf{P} = (m_1 + m_2)\mathbf{V} \; . \tag{2.2}$$

This may be written as

$$\mathbf{V} = \frac{\mathbf{P}}{m_1 + m_2} \tag{2.3}$$

so that the motion of this frame may be given as the rate of change of a preferred coordinate:

$$\mathbf{V} = \frac{d}{dt}\left(\frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2}\right) = \frac{d\mathbf{R}}{dt} \; , \tag{2.4}$$

where $\mathbf{R}$ is the preferred coordinate

$$\mathbf{R} := \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \; . \tag{2.5}$$

Suppose we now choose our coordinates such that $\mathbf{R}$ sits at the origin: $\mathbf{R} \equiv \mathbf{0}$ (we could just as well choose any constant vector $\mathbf{R}_0$). This implies $\mathbf{V} \equiv \mathbf{0}$ and so $\mathbf{P} \equiv \mathbf{0}$. This is the shift to the *center-of-mass coordinate system.* It is defined by the relations

$$m_1\mathbf{r}_1 + m_2\mathbf{r}_2 = \mathbf{0} \; ;$$

$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{0} \; . \tag{2.6}$$

The last equation comes from the definition of the total momentum $\mathbf{P} := \mathbf{p}_1 + \mathbf{p}_2$.

The above relations give six constraints on our twelve-dimensional phase space, bringing the number of degrees of freedom down to six. We should then choose convenient phase space coordinates $(\mathbf{r}, \mathbf{p})$ to describe this subset. The equation for the total momentum shows that choosing $\mathbf{p} := \mathbf{p}_1$ immediately implies

$$\mathbf{p}_1 = \mathbf{p} \; ; \qquad \mathbf{p}_2 = -\mathbf{p} \; . \tag{2.7}$$

The choice of position coordinate follows from more physical considerations: the only direction picked out by the system is along the line extending between the two bodies. Based on this, one chooses

$$\mathbf{r} := \mathbf{r}_1 - \mathbf{r}_2 \; . \tag{2.8}$$

This choice completes the description of the six-dimensional subspace in terms of the coordinates $(\mathbf{r}, \mathbf{p})$. Given a position $\mathbf{r}$, one obtains the positions of the individual particles from the relations

$$\mathbf{r}_1 = \frac{\mu}{m_1}\mathbf{r} \; ; \qquad \mathbf{r}_2 = -\frac{\mu}{m_2}\mathbf{r} \; , \tag{2.9}$$

where

$$\mu := \frac{m_1 m_2}{m_1 + m_2} \tag{2.10}$$

is the reduced mass. Given momentum $\mathbf{p}$, the individual momenta follow from (2.7).

One might ask at this point, how can we transform back to the original full twelve-dimensional system when we have managed to cut it down to six dimensions? Have we not lost information? How can we access the full twelve dimensions when we have only six to work with? The answer lies in the realization that when we initially counted twelve dimensions, we over-estimated. What equations (2.7) and (2.9) are saying is that the motion itself actually only needed six pieces of information. The motion constrains *itself* to six dimensions.

### 2.1.2 Conservation of Angular Momentum

The *conservation of angular momentum* allows us to further simplify the system. Setting

$$\text{const} = \mathbf{L} = \mathbf{r} \times \mathbf{p} \ , \tag{2.11}$$

where $(\mathbf{r}, \mathbf{p})$ are the coordinates obtained from the first simplification, gives three independent constraints on $(\mathbf{r}, \mathbf{p})$. That this is indeed three constraints may be seen from a calculation using the Implicit Function Theorem, which we will discuss later. For now we note that $\mathbf{L}$ is perpendicular to both $\mathbf{r}$ and $\mathbf{p}$, and we may perform a rotation $R_{\mathbf{L}}$ which transforms $\mathbf{L} \mapsto L\hat{\mathbf{z}}$. Under such a transformation, $(\mathbf{r}, \mathbf{p})$ rotates to $(R_{\mathbf{L}}\mathbf{r}, R_{\mathbf{L}}\mathbf{p})$ given by

$$\begin{aligned} (\mathbf{r}, \mathbf{p}) &\mapsto (R_{\mathbf{L}}\mathbf{r}, R_{\mathbf{L}}\mathbf{p}) \\ &= \left( \begin{bmatrix} r\cos\phi \\ r\sin\phi \\ 0 \end{bmatrix} ; \mu\dot{r} \begin{bmatrix} \cos\phi \\ \sin\phi \\ 0 \end{bmatrix} + \mu r\dot{\phi} \begin{bmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{bmatrix} \right) \ , \end{aligned} \tag{2.12}$$

where $r := \|\mathbf{r}\|$ and $\phi$ is the polar angle in the plane. Hence we have new phase coordinates

$$(\tilde{\mathbf{r}}, \tilde{\mathbf{p}}) := (r, \phi; p_r, p_\phi), \tag{2.13}$$

where $(r, \phi)$ are the polar coordinates in the plane perpendicular to $\mathbf{L}$, and $(p_r, p_\phi)$ the corresponding momenta $p_r = \mu\dot{r}$ and $p_\phi = \mu r\dot{\phi}$. The accessible region of the phase space has actually been reduced to, not four, but three dimensions. (This choice of coordinates is not the ideal choice: $\mathbf{L} = \text{const}$ gives *three* constraints on the six-dimensional system, so there is still a relation between the four coordinates we have chosen.)

One more often works with the condition $L = \text{const}$ expressed in the following form:

$$L = \|\mathbf{r} \times \mathbf{p}\| = \mu r^2 \dot{\phi}. \tag{2.14}$$

Writing this constraint in the above form is only possible after the explicit choice of coordinates $(r, \phi; p_r, p_\phi)$. We may recover the coordinates $(\mathbf{r}, \mathbf{p})$ by writing

$$(\mathbf{r}, \mathbf{p}) = R_{\mathbf{L}}^{-1}(\tilde{\mathbf{r}}, \tilde{\mathbf{p}}). \tag{2.15}$$

### 2.1.3 Conservation of Energy and Solution of the System

The *conservation of energy* adds one further constraint to our now two-dimensional system. This reduces the problem to a one-dimensional surface and allows us to solve the system.

The total energy of the original system is given by the Hamiltonian

$$H(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) = \frac{\|\mathbf{p}_1\|^2}{2m_1} + \frac{\|\mathbf{p}_2\|^2}{2m_2} - \frac{Gm_1m_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \; . \tag{2.16}$$

Under the transformation to coordinates $(\mathbf{r}, \mathbf{p})$, this becomes

$$\begin{aligned} H(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) &= \frac{1}{2}\left(\frac{m_2\|\mathbf{p}_1\|^2 + m_1\|\mathbf{p}_2\|^2}{m_1m_2}\right) - \frac{Gm_1m_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \\ &= \frac{\|\mathbf{p}\|^2}{2\mu} - \frac{GM\mu}{\|\mathbf{r}\|} =: H(\mathbf{r}, \mathbf{p}), \end{aligned} \tag{2.17}$$

where $M := m_1 + m_2$ is the total mass. Since rotation does not affect lengths of vectors, we may perform the rotation $R_{\mathbf{L}}$ to write

$$\|\mathbf{p}\|^2 = \|p_r\mathbf{e}_r + p_\phi\mathbf{e}_\phi\|^2 = \mu^2\dot{r}^2 + \mu^2r^2\dot{\phi}^2 \; . \tag{2.18}$$

Using this result and equation (2.14) we may write the Hamiltonian as

$$H(\mathbf{r}, \mathbf{p}) = \frac{\mu\dot{r}^2}{2} + \frac{\mu r^2\dot{\phi}^2}{2} - \frac{GM\mu}{r} = \frac{\mu\dot{r}^2}{2} + \frac{L^2}{2\mu r^2} - \frac{GM\mu}{r} =: H(r, p_r) \; , \tag{2.19}$$

where $p_r = \mu\dot{r}$. This makes explicit the two-dimensional (the radial coordinate and its associated momentum) nature of the system. Imposing the energy conservation constraint

$$H = E, \tag{2.20}$$

where $E$ is a constant, gives a relation between the independent variables and allows one to solve the system. The fact that $\phi$ is cyclic (*i.e.* $H$ does not depend on $\phi$) is the crucial point that allows the constraint $H = E$ to be enough to solve the system. This point will be encountered later in the guise of the isotropy subgroup corresponding to the fixed angular momentum (*cf.* (2.105)).

The solution is given by means of *quadrature*. Applying (2.20) to expression (2.19) yields

$$\frac{dr}{dt} = \sqrt{\frac{2}{\mu}\left(E + \frac{GM\mu}{r} - \frac{L^2}{2\mu r^2}\right)} \ . \tag{2.21}$$

Integration gives

$$t = \int_{r_0}^{r} \frac{\mathrm{d}r'}{\sqrt{\frac{2}{\mu}\left(E + \frac{GM\mu}{r'} - \frac{L^2}{2\mu r'^2}\right)}} \ . \tag{2.22}$$

Performing the integral gives $t = t(r)$, which, upon inversion, provides $r = r(t)$, the equation for the radial motion as a function of time. The angular motion follows from the angular momentum constraint (2.14):

$$\frac{d\phi}{dt} = \frac{L}{\mu r^2} \ . \tag{2.23}$$

Integration gives

$$\phi(t) = \phi_0 + \frac{L}{\mu}\int_0^t \frac{\mathrm{d}t'}{r^2(t')} \ , \tag{2.24}$$

where $r$ is given as a function of time by the inversion of equation (2.22). Thus eqs. (2.22) and (2.24) amount to a complete solution of the system, since the motion of the full two-body system may be obtained under the map

$$t(r) \mapsto r(t) \mapsto \phi(t) \mapsto (r, \phi; p_r, p_\phi) \mapsto (\mathbf{r}, \mathbf{p}) \mapsto (\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2). \tag{2.25}$$

## 2.2 Fundamentals of Symplectic Reduction

In this section we again treat the Kepler Problem, but this time from a more mathematical viewpoint. Specifically, we will use the techniques of symplectic geometry to reformulate the problem.[1]

The first step in our reformulation of the problem is to define a suitable symplectic manifold. That is, we define a manifold $M_0$ with a symplectic structure $\omega_0$, or, more briefly, $(M_0, \omega_0)$. For the two-body problem, the natural phase space is $\mathbb{R}^{12}$ in the guise

$$T^*\mathbb{R}^6 \cong \mathbb{R}^6 \times \mathbb{R}^6 = \{(\mathbf{r}_1, \mathbf{r}_2) \times (\mathbf{p}_1, \mathbf{p}_2)\} = \{(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2)\}, \tag{2.26}$$

with symplectic structure

$$\begin{aligned}\omega_0 = \; & \mathrm{d}p_{1x} \wedge \mathrm{d}r_{1x} + \mathrm{d}p_{1y} \wedge \mathrm{d}r_{1y} + \mathrm{d}p_{1z} \wedge \mathrm{d}r_{1z} \\ & + \mathrm{d}p_{2x} \wedge \mathrm{d}r_{2x} + \mathrm{d}p_{2y} \wedge \mathrm{d}r_{2y} + \mathrm{d}p_{2z} \wedge \mathrm{d}r_{2z}\,.\end{aligned} \tag{2.27}$$

This is written more compactly as

$$\omega_0 = \mathrm{d}\mathbf{p}_1 \overset{\cdot}{\wedge} \mathrm{d}\mathbf{r}_1 + \mathrm{d}\mathbf{p}_2 \overset{\cdot}{\wedge} \mathrm{d}\mathbf{r}_2 \tag{2.28}$$

The notation $\overset{\cdot}{\wedge}$ is read as "wedge-dot" and is a shorthand for the exterior

[1]The following material draws heavily upon notes presented by Stephanie Frank Singer in May 1997 at the Institute for Advanced Study, Princeton, NJ. The material has subsequently been published, *cf.* [51]. See also [1], [4], and [52].

product of column and row vectors of 1-forms:

$$
\begin{aligned}
\mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r} &= \left[\, \mathrm{d}p_x \;\; \mathrm{d}p_y \;\; \mathrm{d}p_z \,\right] \wedge \begin{bmatrix} \mathrm{d}r_x \\ \mathrm{d}r_y \\ \mathrm{d}r_z \end{bmatrix} \\
&= \mathrm{d}p_x \wedge \mathrm{d}r_x + \mathrm{d}p_y \wedge \mathrm{d}r_y + \mathrm{d}p_z \wedge \mathrm{d}r_z . \qquad (2.29)
\end{aligned}
$$

We establish dynamics on $(\mathbb{R}^{12}, \omega_0)$ by defining a Hamiltonian function

$$
\begin{aligned}
H_0 &: \mathbb{R}^{12} \to \mathbb{R} , \\
(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) &\mapsto \frac{\|\mathbf{p}_1\|^2}{2m_1} + \frac{\|\mathbf{p}_2\|^2}{2m_2} - \frac{Gm_1m_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|} . \qquad (2.30)
\end{aligned}
$$

This determines the dynamics by providing the time flow $d/dt$ by means of the relation

$$
\omega_0 \left( \frac{d}{dt}, \cdot \right) = -\,\mathrm{d}H_0(\cdot) , \qquad (2.31)
$$

which in coordinates yields the usual Hamilton's equations:

$$
\frac{d\mathbf{r}_i}{dt} = \frac{\partial H_0}{\partial \mathbf{p}_i} ; \qquad \frac{d\mathbf{p}_i}{dt} = -\frac{\partial H_0}{\partial \mathbf{r}_i} . \qquad (2.32)
$$

Specifically, for the two-body problem specified by the Hamiltonian $H_0$ above, the equations are

$$
\frac{dr_{1j}}{dt} = \frac{p_{1j}}{m_1} ; \qquad \frac{dp_{1j}}{dt} = -\frac{Gm_1m_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3}(r_{1j} - r_{2j}) , \qquad (2.33)
$$

and similarly for $(\mathbf{r}_2, \mathbf{p}_2)$. These are the equations that describe the motion of the system.

There are two things to notice: (1) there are a lot of equations; (2) they are coupled and nonlinear. In short, the system is very difficult to solve in the form given. The task of symplectic reduction is to make these equations "more managable" (*i.e.* fewer, although not necessarily simpler).

### 2.2.1 Symmetry and Symplectic Reduction

What is symplectic reduction? We should get some idea of the broad picture before we actually embark on the process for our specific problem. Simply put, *symplectic reduction* is the procedure by which one exploits symplecto-dynamic[2] symmetries to effectively reduce the orginal system to a smaller symplecto-dynamic system.

What is meant by the term "symplecto-dynamic"? The answer is that we are looking for symmetries of *both* the two-form $\omega_0$ (the *symplectic* structure) *and* the Hamiltonian $H_0$ (the *dynamic* structure). Either one alone is not the whole story from a physics perspective, for only together do we obtain a dynamical structure, *cf.* equation (2.31). We might symbolize this by notating the symplecto-dynamic unit as $(M_0, \omega_0, H_0)$.

How is the procedure accomplished? Suppose we do have a symplecto-dynamic symmetry. Technically what we mean by this is there is some Lie group $G$ that acts on $M_0$. For each $g \in G$, we obtain a map $S_g : M_0 \to M_0$.

[2] Not standard terminology.

This is a symplecto-dynamic symmetry if[3]

$$S_g^*\omega_0 = \omega_0 \qquad \text{and} \qquad H_0 \circ S_g = H_0 \ . \tag{2.34}$$

Now let us restrict our focus to one-parameter subgroups of $G$. In particular, let $\xi$ be an infinitesimal generator of a symplecto-dynamic symmetry; that is, $\xi \in \mathfrak{g}$, where $\mathfrak{g} \cong T_eG$ is the Lie algebra of our symmetry group $G$. Then the one-parameter subgroup of $G$ generated by $\xi$ is

$$G_\xi := \left\{ e^{t\xi} \middle| \, t \in \mathbb{R} \right\} \ . \tag{2.35}$$

Put another way, we could say the one-parameter subgroup $G_\xi$ is the *flow* of $\xi$ in $G$. Since $G$ acts on $M_0$, so does $G_\xi$. Hold on to this idea.

Given our group $G$, we may also consider some function

$$\Phi_0 : M_0 \to \mathfrak{g}^* \ , \tag{2.36}$$

---

[3]The superscript asterisk denotes the *pullback*. Recall that a differentiable map $f : M \to N$ between manifolds defines at a point $\mathbf{x}$ a linear map of tangent spaces $f_{*\mathbf{x}} : T_\mathbf{x}M \to T_{f(\mathbf{x})}N$ as follows. Given $\mathbf{v} \in T_\mathbf{x}M$ and a curve $\phi : \mathbb{R} \to M$ with $\phi(0) = \mathbf{x}$ and $(d\phi/dt)|_{t=0} = \mathbf{v}$, then define $f_{*\mathbf{x}}\mathbf{v}$ by

$$f_{*\mathbf{x}}\mathbf{v} = \left.\frac{d}{dt}\right|_{t=0} f(\phi(t)),$$

*i.e.* $f_{*\mathbf{x}}\mathbf{v}$ is the tangent vector of the curve $f \circ \phi : \mathbb{R} \to N$ at the point $f(\mathbf{x})$. Taking the union of the maps $f_{*\mathbf{x}}$ for all $\mathbf{x} \in M$, we obtain the *differential map* $f_* : TM \to TN$, defined by $f_*\mathbf{v} = f_{*\mathbf{x}}\mathbf{v}$ for $\mathbf{v} \in T_\mathbf{x}M$.

Given a differentiable map $f : M \to N$ between smooth manifolds, and a differential $k$-form $\omega$ on $N$, we define the *pullback* $f^*\omega$ of $\omega$ under $f$ by the relation

$$(f^*\omega)(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \omega(f_*\mathbf{v}_1, \ldots, f_*\mathbf{v}_k),$$

for any tangent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in T_\mathbf{x}M$. The pullback $f^*\omega$ is thus a $k$-form on $M$, whose value on vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is equal to $\omega$'s value on the images of those vectors in $N$.

where $\mathfrak{g}^*$ is the dual of $\mathfrak{g}$. If $\langle \cdot , \cdot \rangle$ is a pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$, then

$$\Phi_0^\xi := \langle \Phi_0(\cdot) , \xi \rangle , \tag{2.37}$$

for fixed $\xi \in \mathfrak{g}$, gives a function from $M_0 \to \mathbb{R}$. Mathematically speaking, this function, inasmuch as it is a function from the manifold to $\mathbb{R}$, may itself be considered as a Hamiltonian on $(M_0, \omega_0)$, *i.e.* we get a symplecto-dynamical system $\left(M_0, \omega_0, \Phi_0^\xi\right)$.[4] Just as in equation (2.31) the Hamiltonian $H_0$ defined the flow $d/dt$, so $\Phi_0^\xi$ defines a Hamiltonian vector field $\xi_{M_0}$ via

$$\omega_0 \left(\xi_{M_0}, \cdot\right) = - \,\mathrm{d}\Phi_0^\xi(\cdot) . \tag{2.38}$$

So $\Phi_0^\xi$ defines a flow on $M_0$ by determining the vector field $\xi_{M_0}$ in the same manner as our usual Hamiltonian $H_0$.

By *coincidence* it might, but need not, happen that the flow of the vector field $\xi_{M_0}$ corresponds to the orbits of $G_\xi$, where $\xi$ is the same vector as in $\Phi_0^\xi$. By "orbits of $G_\xi$" I actually mean orbits of points $m \in M_0$ under the action of $G_\xi$. This is the set

$$G_\xi \cdot m := \{ g_\xi \cdot m | \, g_\xi \in G_\xi \} . \tag{2.39}$$

In this case, where flows of $\xi_{M_0}$ correspond to orbits in $M_0$ under $G_\xi$, then we say *the action of $G_\xi$ on $(M_0, \omega_0)$ is hamiltonian, with hamiltonian function $\Phi_0^\xi$.*

[4]Of course there is no *a priori* reason that $\Phi_0^\xi$ should describe the properties of any physical system. Therefore its choice as a Hamiltonian may be debatable from a physical perspective.

Suppose, even better, that this is the case for *all* $\xi \in \mathfrak{g}$. That is, say *every* one-parameter subgroup $G_\eta$ has a hamiltonian action, with hamiltonian $\Phi_0^\eta$ (where $\Phi_0$ is the *same* function in each case). Then we say *the action of $G$ on $(M_0, \omega_0)$ is hamiltonian, with moment map $\Phi_0$.*[5] Putting all of this together, we have the following definition:

**Definition 1.** $\Phi_0$ *is a* **moment map** *for the $G$-action on* $(M_0, \omega_0)$ *when* both

$$\omega_0\left(\xi_{M_0}, \cdot\right) = -\,\mathrm{d}\Phi_0^\xi(\cdot) \tag{2.40}$$

*and*

$$\xi_{M_0} f(m) = \left.\frac{d}{dt}\right|_{t=0} f\left(e^{t\xi}\cdot m\right) \tag{2.41}$$

*for* all $\xi \in \mathfrak{g}$, *all* $f \in C^1(M_0)$ *and* $m \in M_0$.

This moment map $\Phi_0$ is the main piece of machinery in the process of symplectic reduction. Let's see why this is so.

Let us go back to the beginning of our discussion of symplecto-dynamical symmetry. When a group $G$ gives rise to a symplecto-dynamical symmetry, this says that *appropriate* motions under $G$ are contained *within* the dynamics under the Hamiltonian $H_0$. The idea "contained within" is what the expression

$$H_0 \circ S_g = H_0 \tag{2.42}$$

[5]Notice how it starts to make sense that $\Phi_0$ takes values in $\mathfrak{g}^*$. We want to get a function $M_0 \to \mathbb{R}$, a Hamiltonian, when we input infinitesimal symmetry data, $\xi$. A natural way of obtaining a function from $\xi$ is by using elements of the dual space.

encapsulates. This in essence says that, if you move along under the *dynamics dictated by $G$*, then you are still moving within the dynamics allowed by $H_0$. The statement that the motions under $G$ be "appropriate" is represented by

$$S_g^* \omega_0 = \omega_0 \ . \tag{2.43}$$

The symplectic form $\omega_0$ is the quantity which really relates a function like $H_0$ to the dynamical equations of $H_0$. If the symplectic form is invariant under group motions, then we may say these motions are "dynamical" in the same sense that equation (2.31) says that flows under $d/dt$ are dynamical.

Now $\Phi_0^\xi$ in some sense keeps track of just how "dynamical" the motions of $G$ are. To see this, refer to equations (2.40)-(2.41). The second equation says that flows under $\xi_{M_0}$ are the natural time evolution (motions) under $G$. The first equation says that these motions are *dynamical* in the sense dictated by the symplectic structure $\omega_0$.

### 2.2.2 Group Notions

Let us step back a moment and introduce some new concepts. The way a Lie group $G$ acts can be depicted, via a *representation*,[6] as an action on some vector space. Inasmuch as $\mathfrak{g} \cong T_eG$ is a vector space, then we may depict the action of $G$ by a representation on $\mathfrak{g}$. This is done by means of the adjoint representation.

[6] A *representation* is a homomorphism to linear operators on the target vector space. That is to say, it is a map of the group which preserves the group operation, and whose image points are linear operators on some vector space.

Technically, we may define the *adjoint representation* as follows: if $L_g : G \rightarrow G$ is the left-action $L_g h = gh$, and $R_g : G \rightarrow G$ is the right-action $R_g h = hg$, then we define $\mathrm{Ad}_g : \mathfrak{g} \rightarrow \mathfrak{g}$ by

$$\mathrm{Ad}_g := \left. (L_g R_{g^{-1}})_* \right|_e \ . \tag{2.44}$$

$\mathrm{Ad}_g$ is thus a linear operator on $\mathfrak{g}$, and it preserves the group operation:

$$\mathrm{Ad}_{gh} = \mathrm{Ad}_g \mathrm{Ad}_h \ ; \tag{2.45}$$

*i.e.* it is a group homomorphism.

The important thing to note is that we can now represent the action of a group element $g$ by the action on $\mathfrak{g}$ via $\mathrm{Ad}_g$. In turn, we may also represent the action of $g$ by the action on $\mathfrak{g}^*$ using the *co-adjoint representation*, $\mathrm{Ad}_g^* : \mathfrak{g}^* \rightarrow \mathfrak{g}^*$, given by

$$\langle \, \mathrm{Ad}_g^* \alpha \, , \, \xi \, \rangle = \langle \, \alpha \, , \, \mathrm{Ad}_g \xi \, \rangle. \tag{2.46}$$

Here $\langle \, \cdot \, , \, \cdot \, \rangle$ is the pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$.

Look again at $\Phi_0$. $\Phi_0$ is a map $M_0 \rightarrow \mathfrak{g}^*$ concerned with dynamical motions of the group $G$. $G$ acts on $M_0$. If $\Phi_0$ is going to keep track of these $G$-motions, then we should have a way of representing these motions in $\mathfrak{g}^*$. We do — that is $\mathrm{Ad}^*$, the co-adjoint representation.

One thing is missing, however: $\Phi_0$ keeps track of $G$-motions, not by looking at an element $g \in G$ directly, but rather by looking at $\xi \in \mathfrak{g}$, where[7]

$$g = e^{t\xi} \ . \tag{2.47}$$

[7]This is perhaps misleading phrasing. $G$-motions happen in $M_0$, period. These motions

So rather than useing $\mathrm{Ad}_g^* : \mathfrak{g}^* \to \mathfrak{g}^*$, which for a given $g \in G$ dictates a linear operator on $\mathfrak{g}^*$, instead we shall use $\mathrm{ad}_\xi^* : \mathfrak{g}^* \to \mathfrak{g}^*$ defined by[8]

$$\mathrm{ad}_\xi^* := \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}_{e^{t\xi}}^* \,. \tag{2.48}$$

This, in a sense, gives us a representation of $G$ on $\mathfrak{g}^*$ by looking not at $g \in G$, but at $\xi \in \mathfrak{g}$ for $g = e^{t\xi}$. We will use the terminology *little-co-ad representation* for this construct.

### 2.2.3 The Structure of the Moment Map

Now we have most, if not all, of the concepts we need to make sense of constructions involving the moment map. Given our group $G$ action on $(M_0, \omega_0)$, the existence of a moment map $\Phi_0$ ensures that motions under $G$ are dynamical, in the sense that they obey equations (2.40) and (2.41). It also provides a means of linking $G$-action on $M_0$ to $G$-action on $\mathfrak{g}^*$. It does this by virtue of having the property that the following diagram commutes:

$$\begin{tikzcd}
M_0 \arrow[r, "\Phi_0"] \arrow[d, "S_g"'] & \mathfrak{g}^* \cong T_\xi \mathfrak{g}^* \arrow[d, "\mathrm{ad}_\xi^*"] \\
M_0 \arrow[r, "\Phi_0"] & \mathfrak{g}^* \cong T_\xi \mathfrak{g}^*
\end{tikzcd}$$

affect $\Phi_0$. We may find out how these motions affect $\Phi_0$ either by looking at an element $g$ itself, or by looking at $\xi$ given by $g = e^{t\xi}$. The latter might seem more appropriate if we eventually intend to focus on $\Phi_0^\xi$.

[8]Note that the definition shows that $\mathrm{ad}_\xi^*$ is actually a map from the *tangent space* of $\mathfrak{g}^*$ to itself. Since $\mathfrak{g}^*$ is a vector space, however, its tangent space at a point is naturally isomorphic to itself.

Motion $S_g : M_0 \to M_0$ itself induces motion $\mathrm{ad}_\xi^* : \mathfrak{g}^* \to \mathfrak{g}^*$, where $g = e^{t\xi}$. Commutativity with $\Phi_0$ ensures that the representation of $G$-motion in $\mathfrak{g}^*$ is "dynamically faithful".

Presumably the last bit of the explanation was a little vague. What might now be helpful to solidify some of the concepts is an explanation of the *method* of reduction. (Later we will do the two-body problem in detail, but first the general method.)

The map $\Phi_0 : M_0 \to \mathfrak{g}^*$ takes points in the phase space and associates them with elements of $\mathfrak{g}^*$. The group $G$ acts on $M_0$, as well as on $\mathfrak{g}^*$ (via $\mathrm{Ad}^*$ or $\mathrm{ad}^*$). Each action by an element $g = e^{t\xi}$ in $G$ gives an action in $\mathfrak{g}^*$ by $\mathrm{Ad}_g^*$ (or $\mathrm{ad}_\xi^*$). While $g \cdot m$ might be a different point in $M_0$ from $m \in M_0$, we might still have $\mathrm{Ad}_g^*(\alpha) = \alpha$ (or $\mathrm{ad}_\xi^*(\alpha) = \alpha$) in $\mathfrak{g}^*$. Motion under $G$ in $M_0$ need not imply motion under $G$ in $\mathfrak{g}^*$. Define the *isotropy group at* $\alpha \in \mathfrak{g}^*$ to be

$$\tilde{G}_\alpha := \left\{ g \in G | \, \mathrm{Ad}_g^*(\alpha) = \alpha \right\} \ . \tag{2.49}$$

These are all the $G$-motions in $M_0$ that leave $\alpha \in \mathfrak{g}^*$ fixed.

Now fix $\alpha_0 \in \mathfrak{g}^*$, and look at the level set $\Phi_0^{-1}(\alpha_0) \subseteq M_0$, *i.e.* the set of points $m \in M_0$ for which $\Phi_0(m) = \alpha_0 = \mathrm{const} \in \mathfrak{g}^*$. Many of the points in $\Phi_0^{-1}(\alpha_0)$ may actually be equivalent to one another under $G$-action. *However*, since all points in $\Phi_0^{-1}(\alpha_0)$ are mapped by $\Phi_0$ to the same $\alpha_0 \in \mathfrak{g}^*$, only those $g \in G$ which keep $\alpha_0 \in \mathfrak{g}^*$ fixed may act nontrivially in $\Phi_0^{-1}(\alpha_0)$. This follows from the commutativity of the diagram above. More explicitly, if $\alpha_0 \in \mathfrak{g}^*$ stays fixed when we move along the right leg of the diagram, then commutativity requires that $\alpha_0$ must also stay fixed as we move along the left leg by $S_g$ within $\Phi_0^{-1}(\alpha_0) \subseteq M_0$. These $g$s are then elements of $\tilde{G}_{\alpha_0}$, the isotropy group of $\alpha_0 \in \mathfrak{g}^*$.

If we think of the points in $\Phi_0^{-1}(\alpha_0)$ that differ only by $g \in \tilde{G}_{\alpha_0}$ as (symplecto-) dynamically equivalent, then the quotient space

$$M_1 := \Phi_0^{-1}(\alpha_0)/\tilde{G}_{\alpha_0} \tag{2.50}$$

is the space of symplecto-dynamically *distinct* points in $\Phi_0^{-1}(\alpha_0)$.

What is more, this new manifold $M_1$ inherits a natural symplectic structure $\omega_1$ from $(M_0, \omega_0)$. Technically, if

$$\pi_{\alpha_0} : \Phi_0^{-1}(\alpha_0) \rightarrow M_1 \tag{2.51}$$

is the projection of elements of $\Phi_0^{-1}(\alpha_0)$ onto their equivalence classes; and if

$$\iota_{\alpha_0} : \Phi_0^{-1}(\alpha_0) \rightarrow M_0 \tag{2.52}$$

is the inclusion map, letting one view $\Phi_0^{-1}(\alpha_0)$ as contained in $M_0$; *then*

$$\pi^*_{\alpha_0}\omega_1 = \iota^*_{\alpha_0}\omega_0 \ . \tag{2.53}$$

To put it slightly more intuitively,

$$\pi^*_{\alpha_0}\omega_1 = \omega_0|_{\Phi_0^{-1}(\alpha_0)} \ . \tag{2.54}$$

### 2.2.4 General Method of Symplectic Reduction

We start with a symplecto-dynamical system $(M_0, \omega_0, H_0)$. Along with this we are given a symplecto-dynamical symmetry in the form of a Lie group $G$, which has an associated moment map $\Phi_0$.

When this is the case, we find that we may effectively reduce the original system to a smaller symplecto-dynamical system. To do this, one picks a particular $\alpha_0 \in \mathfrak{g}^*$. In the space $\Phi_0^{-1}(\alpha_0) \subseteq M_0$, this quantity $\alpha_0$ remains invariant under $G$-motions.

Since $\alpha_0$ is a symmetry-induced invariant, and the symmetry motions move "within" the dynamical motions, we may call $\alpha_0$ a *conserved quantity* in $\Phi_0^{-1}(\alpha_0)$. The $g$s that conserve $\alpha_0$ are the elements of the isotropy subgroup

$$\tilde{G}_{\alpha_0} = \left\{ g \in G | \operatorname{Ad}_g^*(\alpha_0) = \alpha_0 \right\}. \tag{2.55}$$

If we put this in the language of infinitesimal symmetries by writing $g = e^{t\xi}$, then we may say $\alpha_0$ is conserved by the group generators in

$$\tilde{G}_{\alpha_0}^{inf} = \left\{ \xi \in \mathfrak{g} | \operatorname{ad}_\xi^*(\alpha_0) = \alpha_0 \right\}. \tag{2.56}$$

This, I believe, is a manner of formulating Emmy Nöther's theorem that every generator of a symmetry Lie group begets a conserved quantity.

Physically speaking, once we find that certain motions of a system allow for conservation of a particular quantity, the specification of a particular value for that conserved quantity relegates these motions to the status of known data which no longer bears directly on the remaining aspects of the dynamical motion. Put slightly more simply, specifying the value of a conserved quantity isolates the motion under the symmetry responsible for the conservation, and it leaves one free to investigate the remaining dynamics.

Mathematically, the passage from $M_0$ to $M_1 := \Phi_0^{-1}(\alpha_0)/\tilde{G}_{\alpha_0}$ is what encapsulates this procedure. Specifying the conserved quantity $\alpha_0$ restricts attention to that part of $M_0$ where the conservation holds. Quotienting by $\tilde{G}_{\alpha_0}$ then removes the data responsible for the conservation, and it leaves only dynamics which is inherently congruent with the conservation law.

To make $M_1$ into a symplecto-dynamical system in its own right, however, it needs to be endowed with a symplecto-dynamic structure. The symplectic structure is given by

$$\pi_{\alpha_0}^* \omega_1 = \omega_0|_{\Phi_0^{-1}(\alpha_0)} \ , \tag{2.57}$$

*i.e.* take the representatives of any class in $\Phi_0^{-1}(\alpha_0)/\tilde{G}_{\alpha_0}$ and evaluate them in $\omega_0$ restricted to $\Phi_0^{-1}(\alpha_0)$.

The dynamic structure $H_1 : M_1 \to \mathbb{R}$ is given by

$$H_1 := \left. H_0 \right|_{M_1} \ . \tag{2.58}$$

That is, $H_1$ is merely $H_0$ written in the coordinates of $M_1 := \Phi_0^{-1}(\alpha_0)/\tilde{G}_{\alpha_0}$.

Through this procedure, we perform a reduction

$$(M_0, \omega_0, H_0) \xrightarrow{\quad \Phi_0,\, G \quad} (M_1, \omega_1, H_1)$$

from a generally larger and more complicated symplecto-dynamical system to a (hopefully) smaller and less complicated system. Once the dynamics of the reduced system are solved completely, one may then translate the solutions back to the original system via $\pi_{\alpha_0}^{-1}$.

Now we can start to work on our specific example.

## 2.3 The Two-Body Problem via Symplectic Reduction

In the preceding section we described the mathematical concepts that will play a role in our subsequent description of the two-body problem. The reason for discussing those mathematical concepts divorced from the physical problem is one of logical clarity: in the process of symplectic reduction there are certain relations that come about merely from the mathematical structures involved, regardless of the physical system being described with those structures. In the last section we have isolated those structures. Specifically, the reader should note that the moment map is a construction associated with a given symplectic manifold and a group action on that manifold. It has nothing *a priori* to do with the physics to which it will be applied. In the groundwork sections (see Appendix A), I have used suggestive notation in calling certain quantities $\mathbf{r}$ and $\mathbf{p}$ and $\mathbf{r} \times \mathbf{p}$, which any physicist worth his salt will immediately imbue with physical meaning. Below we will do just that — imbue these with physical meaning. But the reader should be clear about one fact: up till now, our discussion of group actions and moment maps has been entirely devoid of physics. This hints at the generality of the concepts we have discussed, and at the fact that the above methods should apply to systems other than the two-body problem and the like. That said, we may now stipulate the physical system of our choosing.

As described at the beginning of the mathematical treatment, we start with a symplecto-dynamical system $(M_0, \omega_0, H_0)$, where

$$M_0 = \mathbb{R}^{12} = \{(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2)\}\ , \qquad \omega_0 = \ \mathrm{d}\mathbf{p}_1 \wedge \mathrm{d}\mathbf{r}_1 + \ \mathrm{d}\mathbf{p}_2 \wedge \mathrm{d}\mathbf{r}_2\ , \tag{2.59}$$

and

$$\begin{aligned} H_0 : \mathbb{R}^{12} &\to \mathbb{R}\,, \\ (\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) &\mapsto \frac{1}{2}\left(\frac{\|\mathbf{p}_1\|^2}{m_1} + \frac{\|\mathbf{p}_1\|^2}{m_2}\right) - \frac{Gm_1m_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|}\ . \end{aligned} \tag{2.60}$$

Now finally we have chosen our Hamiltonian based on physical grounds. The phase space $M_0$ is also chosen based on physical considerations. The symplectic structure is the one naturally associated with the manifold. These choices give us the set of equations outlined in (2.33), which determine the evolution of the system. The goal, however, is to recognize at this stage that certain symmetries may be present in the given system and may be exploited to simplify the equations. Mathmatically these symmetries show themselves as group actions, which have associated moment maps, and the moment maps may be employed to reduce the apparent number of degrees of freedom of the system. Specifically, the two-body system above displays translational and rotational symmetries, an observation which may be gleaned from either the mathematical formulation or physical intuition. The procedure, then, is the following:

1. Associate each symmetry with the action of a Lie group: translation with $(\mathbb{R}^3, +)$, rotation with $SO(3)$.

2. With each group, use its associated moment map to express the corresponding conserved quantity: total momentum in the case of $(\mathbb{R}^3, +)$, angular momentum for $SO(3)$.

3. With each group, use level sets of the moment map to focus on dynamics associated with a fixed value of the conserved quantity.

4. In a given level set, quotient by the corresponding isotropy sugroup. This yields a quotient space with effectively fewer degrees of freedom.

5. Choose coordinates on the quotient space.

6. Express the old symplectic structure in terms of these new coordinates.

7. Express the old Hamiltonian in terms of these new coordinates.

Following this procedure for a given group will yield a reduced symplecto-dynamical system, *i.e.* a system with fewer degrees of freedom. The more symmetries, that is, the greater the number of Lie groups acting, the more reductions may be performed and the more the number of degrees of freedom may be reduced. In the present situation, this reduction may be performed twice, finally reducing the system to one degree of freedom and hence integrability.

### 2.3.1 First Reduction

The group $(\mathbb{R}^3, +)$ acts as translations on $\mathbb{R}^{12}$ via the map $S_g : \mathbb{R}^{12} \to \mathbb{R}^{12}$ given by

$$(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) \mapsto (\mathbf{r}_1 + g, \mathbf{r}_2 + g; \mathbf{p}_1, \mathbf{p}_2) \tag{2.61}$$

for $g \in (\mathbb{R}^3, +)$. We must first check $S_g^* \omega_0 = \omega_0$. Computing,

$$S_g^* \omega_0 = \mathrm{d}\mathbf{p}_1 \wedge \mathrm{d}(\mathbf{r}_1 + g) + \mathrm{d}\mathbf{p}_2 \wedge \mathrm{d}(\mathbf{r}_2 + g) = \omega_0 \,, \tag{2.62}$$

since $g$ is constant. We must also check $H_0 \circ S_g = H_0$. This becomes

$$\begin{aligned} H_0 \circ S_g(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) &= H_0(\mathbf{r}_1 + g, \mathbf{r}_2 + g; \mathbf{p}_1, \mathbf{p}_2) \\ &= \frac{1}{2}\left(\frac{\|\mathbf{p}_1\|^2}{m_1} + \frac{\|\mathbf{p}_2\|^2}{m_2}\right) - \frac{G m_1 m_2}{\|(\mathbf{r}_1 + g) - (\mathbf{r}_2 + g)\|} \\ &= H_0(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) \,. \end{aligned} \tag{2.63}$$

Hence $S_g$ is a symplecto-dynamic symmetry. The corresponding moment map is given by

$$\Phi_0 : \mathbb{R}^{12} \to \mathfrak{g}^* \cong \mathbb{R}^3 \,,$$

$$(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) \mapsto \mathbf{p}_1 + \mathbf{p}_2 \,. \tag{2.64}$$

Notice that the image of $\Phi_0$ in $\mathfrak{g}^*$ may be interpreted as the *total momentum* of the system.

Suppose we now fix the total momentum of the system. That is, choose $\alpha_0 \in \mathfrak{g}^* \cong \mathbb{R}^3$. Then look at

$$\Phi_0^{-1}(\alpha_0) = \{\, (\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) |\, \mathbf{p}_1 + \mathbf{p}_2 = \alpha_0 \} \,. \tag{2.65}$$

We then identify points that differ only by translations that keep $\alpha_0$ fixed. To see which translations these are, we need to look at the $\mathrm{Ad}^*$-representation of $(\mathbb{R}^3, +)$. Recall that $(\mathbb{R}^3, +)$ is isomorphic to diagonal matrices with strictly positive entries.

Consider the Ad-representation first. For a group $G$,

$$\mathrm{Ad}_g = \left(g\,(\cdot)\,g^{-1}\right)'\Big|_e \,, \tag{2.66}$$

*cf.* equation (2.44). Noting that we are working with a matrix group, if we describe a path in $G$ by writing $h = e^{t\xi}$, this becomes

$$\mathrm{Ad}_g(\xi) = \frac{d}{dt}\bigg|_{t=0} \left(g e^{t\xi} g^{-1}\right) = g\xi g^{-1}\,, \tag{2.67}$$

for $\xi \in \mathfrak{g}$. In this situation, $\mathrm{Ad}_g^*$ is the usual matrix adjoint of $\mathrm{Ad}_g$, so

$$\mathrm{Ad}_g^* = (\mathrm{Ad}_g)^\dagger = \text{conjugate transpose of } \mathrm{Ad}_g. \tag{2.68}$$

This may be seen from the relation (2.46):

$$\langle\, \mathrm{Ad}_g^*\alpha\,,\, \xi\,\rangle = \langle\, \alpha\,,\, \mathrm{Ad}_g\xi\,\rangle. \tag{2.69}$$

From the above we may note that those $g$s which fix $\xi$ will also fix $\alpha$. Notice that, for

$$g = \begin{pmatrix} e^{t_1} & & \\ & e^{t_2} & \\ & & e^{t_3} \end{pmatrix}, \qquad \text{then} \qquad g^{-1} = \begin{pmatrix} e^{-t_1} & & \\ & e^{-t_2} & \\ & & e^{-t_3} \end{pmatrix}. \tag{2.70}$$

Using the fact that

$$\xi = \begin{pmatrix} \xi_1 & & \\ & \xi_2 & \\ & & \xi_3 \end{pmatrix}, \qquad \text{so that} \qquad e^{t\xi} = \begin{pmatrix} e^{t\xi_1} & & \\ & e^{t\xi_2} & \\ & & e^{t\xi_3} \end{pmatrix}, \tag{2.71}$$

we may compute (2.67):

$$g\xi g^{-1} = \xi. \tag{2.72}$$

Hence $\mathrm{Ad}_g = \mathrm{id}$ and so

$$\langle\, \mathrm{Ad}^*_g \alpha \,,\, \xi \,\rangle = \langle\, \alpha \,,\, \mathrm{Ad}_g \xi \,\rangle = \langle\, \alpha \,,\, \xi \,\rangle \,, \tag{2.73}$$

requiring $\mathrm{Ad}^*_g = \mathrm{id}$ as well. From this we see that, for total momentum $\alpha_0 \in \mathfrak{g}^*$,

$$\mathrm{Ad}^*_g(\alpha_0) = \alpha_0 \quad \text{for } \textit{all } g \in (\mathbb{R}^3, +). \tag{2.74}$$

Thus *all translations fix the total momentum* $\alpha_0$:

$$\tilde{G}_{\alpha_0} = (\mathbb{R}^3, +)\,. \tag{2.75}$$

Returning to (2.65), we want to mod out by all translations $g$ that keep the total momentum fixed. But (2.75) tells us that *all* translations fix $\alpha_0$. Thus we have

$$M_1 := \Phi_0^{-1}(\alpha_0)/\tilde{G}_{\alpha_0} = \Phi_0^{-1}(\alpha_0)/(\mathbb{R}^3, +)\,. \tag{2.76}$$

It is now up to us to choose coordinates on $M_1$. We choose coordinates

$$\begin{cases} \mathbf{r} := \mathbf{r}_1 - \mathbf{r}_2\,, \\ \mathbf{p} := \dfrac{\mu}{m_1}\mathbf{p}_1 - \dfrac{\mu}{m_2}\mathbf{p}_2\,, \end{cases} \qquad \text{where} \qquad \mu := \frac{m_1 m_2}{m_1 + m_2}\;. \tag{2.77}$$

To see that these are good coordinates on $M_1$, first note that, from knowledge of $(\mathbf{r}, \mathbf{p})$ and $\alpha_0$, we may reconstruct $(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2)$ by

$$\begin{cases} \mathbf{p}_1 = \dfrac{\mu}{m_2}\alpha_0 + \mathbf{p}\,, \\ \mathbf{p}_2 = \dfrac{\mu}{m_1}\alpha_0 - \mathbf{p}\,; \end{cases} \qquad \text{and} \qquad \begin{cases} \mathbf{r}_1 = \frac{1}{2}\mathbf{r} + \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2)\,, \\ \mathbf{r}_2 = -\frac{1}{2}\mathbf{r} + \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2)\,. \end{cases} \tag{2.78}$$

This last pair of equations says that

$$\begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix} \quad \text{is equivalent to} \quad \frac{1}{2}\begin{pmatrix} \mathbf{r} \\ -\mathbf{r} \end{pmatrix} \tag{2.79}$$

up to a translation.[9]

Notice that from (2.77) we see that the coordinates $(\mathbf{r}, \mathbf{p})$ are translation invariant; and from (2.78) that

$$\mathbf{p}_1 + \mathbf{p}_2 = \alpha_0 \,, \tag{2.80}$$

as it should. Since this choice of $(\mathbf{r}, \mathbf{p})$ gives good coordinates on $M_1$, we see that

$$M_1 \cong \mathbb{R}^6. \tag{2.81}$$

Now we need to calculate the symplectic structure on $M_1$. This is done by restriction to $M_1$:

$$\begin{aligned}
\omega_1 &= (\,\mathrm{d}\mathbf{p}_1 \wedge \mathrm{d}\mathbf{r}_1 + \,\mathrm{d}\mathbf{p}_2 \wedge \mathrm{d}\mathbf{r}_2)|_{M_1} \\
&= \mathrm{d}\left(\frac{\mu}{m_2}\alpha_0 + \mathbf{p}\right) \wedge \mathrm{d}\left(\frac{1}{2}\mathbf{r}\right) + \mathrm{d}\left(\frac{\mu}{m_1}\alpha_0 - \mathbf{p}\right) \wedge \mathrm{d}\left(-\frac{1}{2}\mathbf{r}\right) \\
&= \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r}\,.
\end{aligned} \tag{2.82}$$

Notice this is the usual symplectic structure on $\mathbb{R}^6$: $\omega_1 = \,\mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r}$.

Finally, we need to calculate the reduced Hamiltonian on $M_1$. We started with the Hamiltonian given in expression (2.60). The reduced Hamiltonian comes from rewriting this in terms of the coordinates $(\mathbf{r}, \mathbf{p})$. Simplified,

---

[9] Viewed another way, translation invariance says that there are really two coordinates, $\mathbf{r}_- := \mathbf{r}_1 - \mathbf{r}_2$ and $\mathbf{r}_+ := \mathbf{r}_1 + \mathbf{r}_2$, and that $\mathbf{r}_+$ is cyclic.

this yields

$$\begin{aligned} H_1 : \mathbb{R}^6 &\to \mathbb{R}\,, \\ (\mathbf{r}, \mathbf{p}) &\mapsto \frac{1}{2}\left(\frac{\|\mathbf{p}\|^2}{\mu} + \frac{\|\alpha_0\|^2}{M}\right) - \frac{GM\mu}{\|\mathbf{r}\|}\,, \end{aligned} \tag{2.83}$$

where $M := m_1 + m_2$, and where $\|\alpha_0\|^2 = (\alpha_0 \mid \alpha_0)$ given by the standard inner product on the Lie algebra of $(\mathbb{R}^3, +)$ is identical to the usual dot-product norm on $\mathbb{R}^3$, *cf.* expression (A.11).

Thus we have performed a reduction from the original symplecto-dynamical system $(M_0, \omega_0, H_0)$ to the new system $(M_1, \omega_1, H_1)$:

$$\begin{tikzcd} \big(\mathbb{R}^{12},\, d\mathbf{p}_1 \wedge d\mathbf{r}_1 + d\mathbf{p}_2 \wedge d\mathbf{r}_2\,, H_0\big) \arrow[d, "\Phi_0 = \mathbf{p}_1 + \mathbf{p}_2", "(\mathbb{R}^3, +)"'] \\ \big(\mathbb{R}^6,\, d\mathbf{p} \wedge d\mathbf{r}\,, H_1\big) \end{tikzcd}$$

with $H_0$ given in eqn. (2.60) and $H_1$ in eqn. (2.83).

Notice the difference between the expression (2.83) and the standard physics version given by (2.17). In the physics version, we moved to the center-of-mass reference frame, where $\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{0}$. By contrast, the term $\|\alpha_0\|^2 / 2M$ in (2.83) keeps track of the value of the total momentum of the original system.

### 2.3.2 Second Reduction

$SO(3)$ acts as rotations on $M_1$ via the map

$$S_g : \mathbb{R}^6 \to \mathbb{R}^6 \,,$$

$$(\mathbf{r}, \mathbf{p}) \mapsto (g\mathbf{r}, g\mathbf{p}) \,, \tag{2.84}$$

for $g \in SO(3)$. We first check that $S_g^* \omega_1 = \omega_1$. This becomes

$$\begin{aligned} S_g^* \omega_1 &= \mathrm{d}\,(g\mathbf{p}) \,\wedge \mathrm{d}\,(g\mathbf{r}) = \left(g^T g\right)\, \mathrm{d}\left(\mathbf{p}^T\right) \wedge \mathrm{d}\mathbf{r} \\ &= \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r} = \omega_1 \;. \end{aligned} \tag{2.85}$$

Also checking $H_1 \circ S_g = H_1$ becomes

$$\begin{aligned} H_1 \circ S_g(\mathbf{r}, \mathbf{p}) &= H_1(g\mathbf{r}, g\mathbf{p}) \\ &= \frac{1}{2}\left(\frac{1}{\mu}\,(g\mathbf{p})^T\,(g\mathbf{p}) + \frac{\|\alpha_0\|^2}{M}\right) - \frac{GM\mu}{\sqrt{(g\mathbf{r})^T\,(g\mathbf{r})}} \\ &= \frac{1}{2}\left(\frac{1}{\mu}\mathbf{p}^T g^T g\mathbf{p} + \frac{\|\alpha_0\|^2}{M}\right) - \frac{GM\mu}{\sqrt{\mathbf{r}^T g^T g\mathbf{r}}} \\ &= \frac{\|\mathbf{p}\|^2}{2\mu} + \frac{\|\alpha_0\|^2}{2M} - \frac{GM\mu}{\|\mathbf{r}\|} = H_1(\mathbf{r}, \mathbf{p}) \,. \end{aligned} \tag{2.86}$$

Hence $S_g$ is a symplecto-dynamic symmetry.

The corresponding moment map is given by

$$\Phi_1 : \mathbb{R}^6 \to \mathfrak{so}(3)^* \,,$$

$$(\mathbf{r}, \mathbf{p}) \mapsto \mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T \;. \tag{2.87}$$

Note that

$$\mathfrak{so}(3)^* \ni \mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T \cong \mathbf{p} \times \mathbf{r} \in (\mathbb{R}^3, \times) \ , \tag{2.88}$$

from equations (A.50) and (A.37). Hence the image of the moment map gives the *angular momentum* of the system.

Now let us fix the total angular momentum of the system. That is, fix $\beta_0^* \in \mathfrak{so}(3)^*$.[10] We then restrict our attention to

$$\Phi_1^{-1}(\beta_0^*) = \left\{(\mathbf{r}, \mathbf{p}) \,\middle|\, \mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T = \beta_0^* \right\} = \left\{(\mathbf{r}, \mathbf{p}) \,\middle|\, \mathbf{p} \times \mathbf{r} = \hat{\beta}_0 \right\} . \tag{2.89}$$

In this notation, $\beta_0 \in \mathfrak{so}(3)$, with $\mathfrak{so}(3)^* \ni \beta_0^* \cong \beta_0 \cong \hat{\beta}_0 \in (\mathbb{R}^3, \times)$. We now want to identify points which differ by rotations that keep $\beta_0^*$ fixed.

Suppose, for simplicity, that

$$\hat{\beta}_0 = \begin{pmatrix} 0 \\ 0 \\ \beta_z \end{pmatrix} , \qquad \text{that is,} \qquad \beta_0 = \begin{pmatrix} 0 & \beta_z & 0 \\ -\beta_z & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{2.90}$$

Why are we concerned with the identification of $\mathfrak{so}(3)^*$ with $\mathfrak{so}(3)$? We want to find those $g \in SO(3)$ that fix $\beta_0^* \in \mathfrak{so}(3)^*$. Since $\mathfrak{so}(3)^* \cong \mathfrak{so}(3)$, with $\beta_0^* \in \mathfrak{so}(3)^*$ identified with $\beta_0 \in \mathfrak{so}(3)$, then we may find the $g$s in $SO(3)$ that fix the representative of $\beta_0^*$ that lies in $\mathfrak{so}(3)$. That is, we may find the $g$s that fix the matrix above. Having done that, the identification $\mathfrak{so}(3)^* \cong \mathfrak{so}(3)$ with $\beta_0^* \cong \beta_0$ tells us that these $g$s will also fix $\beta_0^* \in \mathfrak{so}(3)^*$.

[10] I am using $\beta_0^*$ rather than simply $\beta_0$ because the following arguments will involve switching back and forth between $\mathfrak{so}(3)$ and $\mathfrak{so}(3)^*$. Refer to Section A.8 for the identification between $\mathfrak{so}(3)$ and its dual space.

To figure out which $g$s in $SO(3)$ fix $\beta_0^*$, we go to the adjoint representation and look at $\mathrm{Ad}_g$. In particular, we use expression (2.67), so that the equation we want to solve is[11]

$$g\beta_0 g^{-1} \overset{\heartsuit}{=} \beta_0 \,. \tag{2.91}$$

Remember that here $\mathrm{Ad}_g : \mathfrak{so}(3) \rightarrow \mathfrak{so}(3)$ and $\mathfrak{so}(3) \ni \beta_0 \cong \beta_0^*$.

Now, we want to find *all* $g \in SO(3)$ that satisfy (2.91). If we write out an arbitrary $g$ in terms of sines and cosines, (2.91) will turn into an intractible mess. So instead, we first write $g = e^{t\xi}$, for $\xi \in \mathfrak{so}(3)$. Then (2.91) becomes

$$e^{t\xi}\beta_0 e^{-t\xi} \overset{\heartsuit}{=} \beta_0 \,, \tag{2.92}$$

or, right-multiplying by $g$,

$$e^{t\xi}\beta_0 \overset{\heartsuit}{=} \beta_0 e^{t\xi} \,. \tag{2.93}$$

This equation will hold if $\beta_0$ commutes with $\xi$:

$$[\xi \ , \beta_0] \overset{\heartsuit}{=} 0. \tag{2.94}$$

We may arrive at this result by a slightly different route which may prove instructive. If we write $g = e^{t\xi}$ in its power series form,

$$g = e^{t\xi} = \mathrm{id} + t\xi + \frac{1}{2!}t^2\xi\xi + \dots \,, \tag{2.95}$$

[11] Symbols with a heart above them denote relations which are *intended*, though not yet established. For example, $e^{t\xi}\beta_0 e^{-t\xi} \overset{\heartsuit}{=} \beta_0$ means that we intend (or hope) to show that $e^{t\xi}\beta_0 e^{-t\xi}$ is the same as $\beta_0$, though it has not yet been proven.

then substitution into (2.93) gives

$$\left(\mathrm{id} + t\xi + \frac{1}{2!}t^2\xi\xi + \dots\right)\beta_0 \overset{\heartsuit}{=} \beta_0\left(\mathrm{id} + t\xi + \frac{1}{2!}t^2\xi\xi + \dots\right). \tag{2.96}$$

If we now differentiate both sides at $t = 0$ (*i.e.* differentiate at the identity), we get

$$\xi\beta_0 \overset{\heartsuit}{=} \beta_0\xi\,, \qquad \text{that is,} \qquad [\xi\ ,\ \beta_0] \overset{\heartsuit}{=} 0. \tag{2.97}$$

This is the passage to the *little-ad representation*.

So we need to solve (2.94). Since $g$ was an arbitrary rotation, the $\xi$ will be an arbitrary infinitesimal rotation:

$$\xi = \begin{pmatrix} 0 & \xi_z & -\xi_y \\ -\xi_z & 0 & \xi_x \\ \xi_y & -\xi_x & 0 \end{pmatrix}. \tag{2.98}$$

Then (2.94) becomes

$$\begin{pmatrix} 0 & 0 & -\beta_z\xi_x \\ 0 & 0 & -\beta_z\xi_y \\ \beta_z\xi_x & \beta_z\xi_y & 0 \end{pmatrix} \overset{\heartsuit}{=} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.99}$$

This gives[12]

$$\xi_x = 0 \qquad \text{and} \qquad \xi_y = 0, \tag{2.100}$$

so that

$$g = \exp\left[t\begin{pmatrix} 0 & \xi_z & 0 \\ -\xi_z & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right]. \tag{2.101}$$

[12]This is for the case $\hat{\beta}_0 \neq \mathbf{0}$. The degenerate case of no angular momentum is one in which motion occurs only along a line.

This is a rotation around the $z$-axis. Hence the $g$s in $SO(3)$ that leave $\beta_0$, and so $\beta_0^*$ as well, fixed are rotations about the $z$-axis. This fits with our intuition from the form of $\hat{\beta}_0$ in (2.90).

All this leads us to conclude

$$\begin{aligned}\tilde{G}_{\beta_0^*} &= \left\{ g \in SO(3) \middle| g = \exp\left[ t \begin{pmatrix} 0 & \xi_z & 0 \\ -\xi_z & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] \right\} \\ &= \{\text{rotations about the } z\text{-axis}\}\,. \end{aligned} \tag{2.102}$$

Then we may define

$$M_2 := \Phi_1^{-1}\left(\beta_0^*\right)/\tilde{G}_{\beta_0^*} = \Phi_1^{-1}\left(\beta_0^*\right)/\left\{\text{rotations about the } z\text{-axis}\right\}. \tag{2.103}$$

Remember this is for the particular choice of $\beta_0^*$ with $\beta_0$ as in equation (2.90).

We must now choose coordinates on $M_2$. This requires we know the dimension of $M_2$, which in turn requires we know the dimension of $\Phi_1^{-1}\left(\beta_0^*\right)$. A purely mathematical approach making use of the *Implicit Function Theorem* shows that $\dim \Phi_1^{-1}\left(\beta_0^*\right) = 3$. (See Appendix B.) To figure out the dimension of $M_2$, however, we need not only the dimension of $\Phi_1^{-1}\left(\beta_0^*\right)$, but also the dimension of $\tilde{G}_{\beta_0^*}$ — the isotropy subgroup corresponding to $\beta_0^*$. Looking back at (2.102) and viewing $\tilde{G}_{\beta_0^*}$ as a manifold, we see that $\dim \tilde{G}_{\beta_0^*} = 1$, since the map $\psi : \tilde{G}_{\beta_0^*} \to \mathbb{R}$ given by

$$\exp\left[ t \begin{pmatrix} 0 & \xi_z & 0 \\ -\xi_z & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] \mapsto t \tag{2.104}$$

is a good coordinate chart. Any point in $\tilde{G}_{\beta_0^*}$ can be reached simply by specifying the value of the parameter $t$.

Finally, when we quotient $\Phi_1^{-1}(\beta_0^*)$ with $\tilde{G}_{\beta_0^*}$, we reduce the dimension of the former by the dimension of the latter. Hence

$$\dim\left(\Phi_1^{-1}(\beta_0^*)/\tilde{G}_{\beta_0^*}\right) = \dim\Phi_1^{-1}(\beta_0^*) - \dim\tilde{G}_{\beta_0^*} = 2. \tag{2.105}$$

Be sure to understand the meaning of this. This says that the dimension of $\Phi_1^{-1}(\beta_0^*) \subseteq \{(\mathbf{r},\mathbf{p})\}$, when quotiented by $\tilde{G}_{\beta_0^*}$, is 2. That is, we have reduced the dimension of the *total phase space* to 2. This means we have left only *one coordinate and one momentum*. This is a *one degree-of-freedom system*.

Now we must choose coordinates on $M_2 = \Phi_1^{-1}(\beta_0^*)/\tilde{G}_{\beta_0^*}$. Since we only have a two-dimensional phase space remaining, we must choose one physical coordinate and its associated momentum. With that in mind, the following coordinate choice may seem logical:

$$r := \|\mathbf{r}\|, \quad p_r := \frac{\mathbf{p}\cdot\mathbf{r}}{\|\mathbf{r}\|}. \tag{2.106}$$

We should verify that we can recover $(\mathbf{r},\mathbf{p})$ from knowledge of $(r, p_r)$.

**Claim 2.** *The sextuplet*

$$\left(\begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} p_r \\ -\beta_z/r \\ 0 \end{bmatrix}\right) \tag{2.107}$$

*is equivalent to* $(\mathbf{r},\mathbf{p})$ *under the diagonal* $SO(3)$ *action.*[13]

[13]More importantly, the pairs of coordinates are equivalent under the action of the *isotropy subgroup* $\tilde{G}_{\beta_0^*}$ in particular.

*Proof :* Recall we have chosen $\beta_0^*$ such that

$$\hat{\beta}_0 = \begin{bmatrix} 0 \\ 0 \\ \beta_z \end{bmatrix} = \beta_z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{2.108}$$

We want to show there is a unique $g_0 \in SO(3)$ which fixes the vector $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ and rotates $\mathbf{r}$ to $\begin{bmatrix} r & 0 & 0 \end{bmatrix}^T$.

Note that $\mathbf{r}$ is perpendicular to $\hat{\beta}_0$ (since $\mathbf{p} \times \mathbf{r} = \hat{\beta}_0$), and has length $r$. Hence there is some angle $\phi$ such that

$$\mathbf{r} = \begin{bmatrix} r\cos\phi \\ r\sin\phi \\ 0 \end{bmatrix}. \tag{2.109}$$

Now choose

$$g_0 := \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{so that} \quad g_0^{-1} \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} = \mathbf{r}\,. \tag{2.110}$$

Notice that $g_0$ is a rotation about the $z$-axis, and so fixes $\hat{\beta}_0$ (*i.e.* $g_0 \in \tilde{G}_{\beta_0^*}$ as we would expect, since this is the data we are removing in the process of modding out).

We now need to show

$$\mathbf{p} \overset{\heartsuit}{=} g_0^{-1} \begin{bmatrix} r \\ -\beta_z/r \\ 0 \end{bmatrix}. \tag{2.111}$$

We first express $\mathbf{p}$ in terms of two mutually perpendicular vectors in the plane perpendicular to $\hat{\beta}_0$, specifically

$$\frac{\mathbf{r}}{\|\mathbf{r}\|} \quad \text{and} \quad \frac{\mathbf{r} \times \hat{\beta}_0}{\|\mathbf{r}\| \left\|\hat{\beta}_0\right\|}. \tag{2.112}$$

Projecting $\mathbf{p}$ onto these vectors, we have

$$\mathbf{p} = \left(\mathbf{p} \cdot \frac{\mathbf{r}}{r}\right) \frac{\mathbf{r}}{r} + \left(\mathbf{p} \cdot \frac{\mathbf{r} \times \hat{\beta}_0}{r \left\|\hat{\beta}_0\right\|}\right) \frac{\mathbf{r} \times \hat{\beta}_0}{r \left\|\hat{\beta}_0\right\|}$$

$$= p_r\, g_0^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \frac{\left\|\hat{\beta}_0\right\|}{r} \frac{1}{r \left\|\hat{\beta}_0\right\|} \left( g_0^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right) \times \hat{\beta}_0 \,. \tag{2.113}$$

In the second equality we have used the fact that

$$\mathbf{p} \cdot \left(\mathbf{r} \times \hat{\beta}_0\right) = \hat{\beta}_0 \cdot (\mathbf{p} \times \mathbf{r}) = \hat{\beta}_0 \cdot \hat{\beta}_0 = \left\|\hat{\beta}_0\right\|^2, \tag{2.114}$$

since we are on the level set with $\Phi_1(\mathbf{r}, \mathbf{p}) = \mathbf{p} \times \mathbf{r} = \hat{\beta}_0$. Furthermore, since $g_0$ fixes $\hat{\beta}_0$, so does $g_0^{-1}$, and hence we have

$$\mathbf{p} = p_r\, g_0^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \frac{1}{r^2} \left( g_0^{-1} \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} \right) \times \left( g_0^{-1} \hat{\beta}_0 \right)$$

$$= p_r\, g_0^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \frac{1}{r^2}\, g_0^{-1} \left( \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ 0 \\ \beta_z \end{bmatrix} \right). \tag{2.115}$$

Performing the cross product,

$$\mathbf{p} = g_0^{-1} \begin{bmatrix} p_r \\ -\beta_z/r \\ 0 \end{bmatrix}. \tag{2.116}$$

Hence

$$g_0 \mathbf{r} = \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad g_0 \mathbf{p} = \begin{bmatrix} p_r \\ -\beta_z/r \\ 0 \end{bmatrix}, \tag{2.117}$$

so that we have the equivalence

$$\left( \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} p_r \\ -\beta_z/r \\ 0 \end{bmatrix} \right) \sim (\mathbf{r}, \mathbf{p}) \tag{2.118}$$

under the rotation $g_0 \in \tilde{G}_{\beta_0^*}$. $\square$

This establishes that $(r, p_r)$ are good coordinates on $M_2$. Noting that $r$ can take only strictly positive values, while $p_r$ can take arbitrary real values, we have

$$M_2 \cong \mathbb{R}^{>0} \times \mathbb{R}, \quad \text{where} \quad \mathbb{R}^{>0} := \{r \in \mathbb{R} |\, r > 0\}\,. \tag{2.119}$$

We may also compute the symplectic structure $\omega_2$ on $M_2$. We have

$$\omega_2 = \mathrm{d}\begin{bmatrix} p_r \\ -\beta_z/r \\ 0 \end{bmatrix} \curlywedge \mathrm{d}\begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} = \mathrm{d}p_r \wedge \mathrm{d}r\,. \tag{2.120}$$

This is the usual symplectic structure on $\mathbb{R}^{>0} \times \mathbb{R}$: $\omega_2 = \mathrm{d}p_r \wedge \mathrm{d}r$.

Finally, we are left to calculate the reduced Hamiltonian on $M_2$. From the first reduction, we obtained the first reduced Hamiltonian in equation (2.83). Further reducing this, we substitute the new coordinates for the old:

$$\begin{aligned} H_2(r, p_r) = H_1(g_0\mathbf{r}, g_0\mathbf{p}) &= \frac{1}{2\mu}\left\| \begin{bmatrix} p_r \\ -\beta_z/r \\ 0 \end{bmatrix} \right\|^2 + \frac{\|\alpha_0\|^2}{2M} - GM\mu \left\| \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} \right\|^{-1} \\ &= \frac{p_r^2}{2\mu} + \frac{\beta_z^2}{2\mu r^2} + \frac{\|\alpha_0\|^2}{2M} - \frac{GM\mu}{r}\,. \end{aligned} \tag{2.121}$$

Thus we obtain a dynamical structure

$$\begin{aligned} &H_2 : \mathbb{R}^{>0} \times \mathbb{R} \to \mathbb{R}\,; \\ &\qquad (r, p_r) \mapsto \frac{p_r^2}{2\mu} + \left( \frac{\beta_z^2}{2\mu r^2} + \frac{\|\alpha_0\|^2}{2M} - \frac{GM\mu}{r} \right)\,. \end{aligned} \tag{2.122}$$

This completes the second reduction

$$
\begin{tikzcd}
\left(\mathbb{R}^6,\, \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r}\,, H_1\right) \arrow[d, "SO(3)"', "\Phi_1 = \mathbf{p} \times \mathbf{r}"] \\
\left(\mathbb{R}^{>0} \times \mathbb{R},\, \mathrm{d}p_r \wedge\, \mathrm{d}r\,, H_2\right)
\end{tikzcd}
$$

with $H_1$ given in eqn. (2.83) and $H_2$ in eqn. (2.122).

### 2.3.3 Reduction Summary

In the paragraph preceding Section 2.2.1 I stated that our task is to make equations (2.33) — the twelve Hamilton's equations for the two-body problem — "more managable" using the method of symplectic reduction. We have now performed a series of reductions:

$$
\begin{tikzcd}
\left(\mathbb{R}^{12},\, \mathrm{d}\mathbf{p}_1 \wedge \mathrm{d}\mathbf{r}_1 +\, \mathrm{d}\mathbf{p}_2 \wedge \mathrm{d}\mathbf{r}_2\,, H_0\right) \arrow[d, "G_0 = (\mathbb{R}^3, +)"', "\Phi_0 = \mathbf{p}_1 + \mathbf{p}_2"] \\
\left(\mathbb{R}^6,\, \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r}\,, H_1\right) \arrow[d, "G_1 = SO(3)"', "\Phi_1 = \mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T"] \\
\left(\mathbb{R}^{>0} \times \mathbb{R},\, \mathrm{d}p_r \wedge\, \mathrm{d}r\,, H_2\right).
\end{tikzcd}
$$

In the end, have we made the equations more managable? Let us calculate Hamilton's equations in $(M_2, \omega_2, H_2)$, where

$$M_2 \cong \mathbb{R}^{>0} \times \mathbb{R};$$

$$\omega_2 = \mathrm{d}p_r \wedge \mathrm{d}r;$$

$$H_2(r, p_r) = \frac{p_r^2}{2\mu} + \left(\frac{\beta_z^2}{2\mu r^2} + \frac{\|\alpha_0\|^2}{2M} - \frac{GM\mu}{r}\right) . \tag{2.123}$$

We do this by going back to equation (2.31), written in terms of the reduced system:

$$\omega_2\left(\frac{d}{dt}, \cdot\right) = -\,\mathrm{d}H_2(\cdot) . \tag{2.124}$$

Substituting the explicit form of $H_2$ from (2.123), we get

$$\frac{dr}{dt} = \frac{p_r}{\mu}, \qquad \frac{dp_r}{dt} = \frac{\beta_z^2}{2\mu r^3} - \frac{GM\mu}{r^2} . \tag{2.125}$$

Thus we get *two* equations for the motion, rather than the *twelve* equations (2.33). At least in this sense, the reduced system seems "more managable". More importantly for the case at hand, we have reduced to a one-degree-of-freedom system, which is therefore integrable.

### 2.3.4 Epilogue to the Two-Body Problem

Now that we have approached the two-body problem from both the physical and the mathematical perspectives, we might take a moment to see what insights we may have gained from contrasting the viewpoints.

Return for a moment to equations (2.64)-(2.65), where we fixed the total momentum of the system:

$$\mathbf{p}_1 + \mathbf{p}_2 = \alpha_0. \tag{2.126}$$

We found that, after the first reduction to $M_1$ was completed, we obtained a new Hamiltonian

$$H_1(\mathbf{r}, \mathbf{p}) = \frac{\|\mathbf{p}\|^2}{2\mu} + \left( \frac{\|\alpha_0\|^2}{2M} - \frac{GM\mu}{\|\mathbf{r}\|} \right). \tag{2.127}$$

In this case we may view the term $\|\alpha_0\|^2 /2M$ as a shift of the zero of the potential[14] or of the total energy itself.

Alternatively, realizing that momentum is a frame-dependent quantity, we may take the term $\|\alpha_0\|^2 /2M$ as keeping track of which reference frame we are in. That is, $\alpha_0$, via equation (2.126), tells us which frame we are in by telling us what the total momentum of that frame is. In the case $\alpha_0 = 0$, we return to the *center-of-mass frame*.

In the usual mathematical treatments of this topic — symplectic reduction and the two-body problem — there is frequently a remark to the effect that we have kept some extra information with $\alpha_0$ that "physicists" typically throw away. It is not clear, however, that the retaining of $\alpha_0$ really gives any further insight. The *physics* of the problem is frame-independent. Speaking more mathematically, it is not the Hamiltonian $H_1$, but the Hamiltonian *and*

[14] *cf.* (2.83) and the remarks that follow.

Hamilton's equations that describe the physical situation. By the very nature of Hamilton's equations, any such constant has no physical significance.

# Chapter 3

# de motu duorum corporum vi Darvini se mutuo attrahentium

In 1920, C.G. Darwin published work on the motion of two charged bodies [9]. Inspired by the work of Bohr and Sommerfeld, Darwin derived both the Lagrangian and the Hamiltonian for a system of two charged particles in relativistic motion, valid up to terms of order $(v/c)^2$ inclusive. This approximation has the virtue of replacing the problem of retarded times with action-at-a-distance corrections to the Kepler problem. Darwin solved the system and applied his results to the spectrum of the hydrogen atom.

Since Darwin's original paper, the so-called Darwin Lagrangian has remained for the most part a tool for quantum mechanical studies [2, 6, 31]. Some[1] have advocated studying the Darwin Lagrangian as a classical system in its own right. Few [11, 12, 22] have taken on such studies, and the present investigation falls into this category.

This chapter focuses on the Darwin Two-Body Problem (D2BP). In particular, attention is paid to how symplectic reduction applies to the system.

[1]Perhaps the most notable among these is the paper by Kaufman and Rostler [32].

We then proceed to find the quadrature for the system.[2] Finally we investigate the existence of fixed points of the equations of motion, and the physicality of the conditions for their existence.

## 3.1 Introduction

We begin with a derivation of the Darwin Lagrangian. The derivation given here follows Landau[3] The starting point is the relativistic Lagrangian for a charge $e_a$ in a *given* external field:

$$L_a = -m_a c^2 \sqrt{1 - \frac{v_a^2}{c^2}} - e_a \phi + \frac{e_a}{c} \mathbf{A} \cdot \mathbf{v}_a \ . \tag{3.1}$$

Choosing the Lorenz gauge $\nabla \cdot \mathbf{A} + (1/c)(\partial\phi/\partial t) = 0$ so that the equations for $\phi$ and $\mathbf{A}$ decouple, one obtains the retarded potentials

$$\phi = \int \frac{\rho_{t-R/c}}{R} \, \mathrm{d}^3\mathbf{x}';$$

$$\mathbf{A} = \frac{1}{c} \int \frac{\mathbf{j}_{t-R/c}}{R} \, \mathrm{d}^3\mathbf{x}' \ , \tag{3.2}$$

with $\rho_{t-R/c}$ and $\mathbf{j}_{t-R/c}$ the charge density and charge current density, respectively, at the retarded time. $R$ is the distance from the integration point $\mathbf{x}'$ to the field observation point $\mathbf{x}$.

The next step is to expand in the time slot $t' = t - R/c$. Up to terms

[2] As mentioned above, Darwin himself already found a solution for the system, but by different means.

[3] [33], §65. For a slightly different approach, see [30], §12.7.

of second order in $v/c$, the electric potential becomes

$$\begin{aligned}
\phi &= \int \frac{\rho_t}{R}\,\mathrm{d}^3\mathbf{x}' - \frac{\partial}{\partial t}\int \left(\frac{R}{c}\right)\frac{\rho_t}{R}\,\mathrm{d}^3\mathbf{x}' + \frac{\partial^2}{\partial t^2}\int \left(\frac{R}{c}\right)^2\frac{\rho_t}{R}\,\mathrm{d}^3\mathbf{x}' \\
&= \int \frac{\rho_t}{R}\,\mathrm{d}^3\mathbf{x}' - \frac{1}{c}\frac{\partial}{\partial t}\int \rho_t\,\mathrm{d}^3\mathbf{x}' + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\int R\rho_t\,\mathrm{d}^3\mathbf{x}' \\
&= \int \frac{\rho_t}{R}\,\mathrm{d}^3\mathbf{x}' + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\int R\rho_t\,\mathrm{d}^3\mathbf{x}'\ ,
\end{aligned} \tag{3.3}$$

where the last step follows by charge conservation,

$$\frac{\partial}{\partial t}\int \rho_t\,\mathrm{d}^3\mathbf{x}' = 0\ , \tag{3.4}$$

since the total charge in the volume is given by the above integral and it remains constant in time. The vector potential is simpler: since it is everywhere in the Lagrangian multiplied by $1/c$, we need only keep the first term

$$\mathbf{A} = \frac{1}{c}\int \frac{\mathbf{j}_t}{R}\,\mathrm{d}^3\mathbf{x}' = \frac{1}{c}\int \frac{\rho_t\mathbf{v}}{R}\,\mathrm{d}^3\mathbf{x}' \tag{3.5}$$

to obtain terms of order $(v/c)^2$.

At the moment, we're dealing with the Lagrangian for a single particle. This means that the density $\rho$ is a single delta function, so that the integral expressions for $\phi$ and $\mathbf{A}$ collapse about the point:

$$\begin{aligned}
\phi &= \int \frac{\rho}{R}\,\mathrm{d}^3\mathbf{x}' + \frac{1}{2c^2}\frac{\partial^2}{\partial t^2}\int R\rho\,\mathrm{d}^3\mathbf{x}' = \frac{e}{R} + \frac{e}{2c^2}\frac{\partial^2 R}{\partial t^2}\ ; \\
\mathbf{A} &= \frac{e\mathbf{v}}{cR}\ .
\end{aligned} \tag{3.6}$$

The expression for $\phi$ contains a term with two time derivatives. This may be reduced to one time derivative by a gauge transformation. Since a transformation of the form

$$\phi' = \phi - \frac{1}{c}\frac{\partial f}{\partial t} \ ,$$

$$\mathbf{A}' = \mathbf{A} + \nabla f \tag{3.7}$$

leaves the physics invariant, then comparison of the first equation above with the expression derived for $\phi$ suggests we set

$$\frac{1}{c}\frac{\partial f}{\partial t} = \frac{e}{2c^2}\frac{\partial^2 R}{\partial t^2} \ . \tag{3.8}$$

That is,

$$f := \frac{e}{2c^2}\frac{\partial R}{\partial t} \ . \tag{3.9}$$

This leaves us with new potentials

$$\phi' = \frac{e}{R} \ ,$$

$$\mathbf{A}' = \frac{e\mathbf{v}}{cR} + \frac{e}{2c}\nabla\frac{\partial R}{\partial t} = \frac{e\mathbf{v}}{cR} + \frac{e}{2c}\frac{\partial}{\partial t}\nabla R \ . \tag{3.10}$$

It now remains to compute the last term in the expression for $\mathbf{A}'$. We need to first be a little more explicit about the variables chosen. If the charge $e$ is placed at a point $\mathbf{y}$ and we observe the field $\mathbf{A}'$ at a point $\mathbf{x}$ at time $t$, then the correct expression for the vector potential is

$$\mathbf{A}'(\mathbf{x}, t) = \frac{e\mathbf{v}}{c\,\|\mathbf{x} - \mathbf{y}(t)\|} + \frac{e}{2c}\frac{\partial}{\partial t}\nabla\,\|\mathbf{x} - \mathbf{y}(t)\| \ , \tag{3.11}$$

where

$$\mathbf{v} = \frac{d\mathbf{y}(t)}{dt} \ . \tag{3.12}$$

The gradient, however, is taken with respect to the coordinates of $\mathbf{x}$. So we have the following:

$$\nabla_{\mathbf{x}} \|\mathbf{x}-\mathbf{y}(t)\| = \frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|} = \hat{\mathbf{r}} \tag{3.13}$$

and also

$$\frac{\partial}{\partial t}\left(\frac{1}{\|\mathbf{x}-\mathbf{y}\|}\right) = \frac{-1}{R^2}\frac{1/2}{R}\, 2\mathbf{R}\cdot(-\mathbf{v}) = \frac{1}{R^2}\hat{\mathbf{r}}\cdot\mathbf{v} \tag{3.14}$$

using the fact that $\mathbf{v} = d\mathbf{y}/dt$. This gives

$$\begin{aligned}\frac{\partial\hat{\mathbf{r}}}{\partial t} = \frac{\partial}{\partial t}\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|} &= \frac{1}{\|\mathbf{x}-\mathbf{y}\|}(-\mathbf{v}) + (\mathbf{x}-\mathbf{y}(t))\frac{1}{R^2}\hat{\mathbf{r}}\cdot\mathbf{v} \\ &= \frac{1}{\|\mathbf{x}-\mathbf{y}\|}\left[-\mathbf{v}+\hat{\mathbf{r}}(\hat{\mathbf{r}}\cdot\mathbf{v})\right] \ . \end{aligned}\tag{3.15}$$

Substituting back into $\mathbf{A}'$, we obtain the final expressions for the potentials

$$\begin{aligned}\mathbf{A}'(\mathbf{x},t) &= \frac{e\mathbf{v}}{c\,\|\mathbf{x}-\mathbf{y}(t)\|} + \frac{e}{2c}\frac{1}{\|\mathbf{x}-\mathbf{y}(t)\|}\left[-\mathbf{v}+\hat{\mathbf{r}}(\hat{\mathbf{r}}\cdot\mathbf{v})\right] \\ &= \frac{e[\mathbf{v}+(\mathbf{v}\cdot\hat{\mathbf{r}})\hat{\mathbf{r}}]}{2c\,\|\mathbf{x}-\mathbf{y}(t)\|}\ ; \\ \phi'(\mathbf{x},t) &= \frac{e}{\|\mathbf{x}-\mathbf{y}(t)\|}\end{aligned}\tag{3.16}$$

along with the definition given in (3.12).

Returning to the Lagrangian $L_a$ we may use the expansion

$$\sqrt{1-x} \approx 1 - \frac{x}{2} - \frac{x^2}{4} \tag{3.17}$$

for small $x$ to write

$$-m_a c^2\sqrt{1-\frac{v_a^2}{c^2}} \approx -m_a c^2\left[1-\frac{1}{2}\frac{v_a^2}{c^2}-\frac{1}{4}\left(\frac{v_a^2}{c^2}\right)^2\right] = -m_a c^2 + \frac{m_a v_a^2}{2} + \frac{m_a v_a^4}{8c^2}\ . \tag{3.18}$$

We may drop the term $-m_a c^2$ as a constant which does not contribute to the equations of motion. Thus, using the above equation and the expressions (3.16) for $\phi'$ and $\mathbf{A}'$, the Lagrangian for a particle $a$ interacting with other particles $b$ becomes

$$L_a = \frac{m_a v_a^2}{2} + \frac{m_a v_a^4}{8c^2} - e_a \sum_{b \neq a} \frac{e_b}{R_{ab}} + \frac{e_a}{2c^2} \sum_{b \neq a} \frac{e_b}{R_{ab}} \left[ \mathbf{v}_a \cdot \mathbf{v}_b + (\mathbf{v}_a \cdot \hat{\mathbf{r}}_{ab}) (\mathbf{v}_b \cdot \hat{\mathbf{r}}_{ab}) \right] , \tag{3.19}$$

where $\sum_{b \neq a}$ denotes summation over the index $b$ *ex*-cluding the charge $a$, and $R_{ab}$ is the distance between particles $a$ and $b$.

From here it is straightforward to obtain the corresponding Hamiltonian. First note in the first and third terms of (3.19) the Lagrangian for the Kepler problem, $L_0$. The corresponding Hamiltonian is, of course,

$$H_0 = \frac{p_a}{2m_a} + e_a \sum_{b \neq a} \frac{e_b}{R_{ab}}. \tag{3.20}$$

The remaining terms in (3.19) are small corrections to $L_0$. Basic mechanics shows that small changes to $L$ and $H$ are equal in magnitude and opposite in sign. (See Appendix C.) Thus, switching the signs of the remaining terms in $L_a$ and adding them to $H_0$, with $\mathbf{v}_a = \mathbf{p}_a / m_a$, we find

$$\begin{aligned} H_a = &\frac{p_a}{2m_a} - \frac{p_a^4}{8c^2 m_a^3} + e_a \sum_{b \neq a} \frac{e_b}{R_{ab}} \\ &- \frac{e_a}{2c^2} \sum_{b \neq a} \frac{e_b}{2c^2 m_a m_b R_{ab}} \left[ \mathbf{p}_a \cdot \mathbf{p}_b + (\mathbf{p}_a \cdot \hat{\mathbf{r}}_{ab}) (\mathbf{p}_b \cdot \hat{\mathbf{r}}_{ab}) \right] . \end{aligned} \tag{3.21}$$

Simplifying to the case of two charged bodies:

$$H(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) = \frac{\|\mathbf{p}_1\|^2}{2m_1} + \frac{\|\mathbf{p}_2\|^2}{2m_2} - \frac{\|\mathbf{p}_1\|^4}{8c^2{m_1}^3} - \frac{\|\mathbf{p}_2\|^4}{8c^2{m_2}^3} + \frac{e_1 e_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|}$$
$$-\frac{e_1 e_2}{2c^2 m_1 m_2} \frac{1}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \left[\mathbf{p}_1 \cdot \mathbf{p}_2 + (\mathbf{p}_1 \cdot \hat{\mathbf{r}})(\mathbf{p}_2 \cdot \hat{\mathbf{r}})\right] , \tag{3.22}$$

where $\hat{\mathbf{r}}$ is the unit vector along the straight line joining the two bodies.

As with the standard (Newtonian) two-body problem (2BP or N2BP), one may apply symplectic reduction to the Darwin two-body problem. This merely formalizes the usual physics methods of switching to the center-of-mass reference frame and using rotational invariance.

### 3.1.1 Translation Invariance

The mathematical formalization of transferring to the center-of-mass frame is reducing under the action of the group $(\mathbb{R}^3, +)$. As in the N2BP, we have $(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) \mapsto (\mathbf{r}_1 + g, \mathbf{r}_2 + g; \mathbf{p}_1, \mathbf{p}_2)$. Since

$$\hat{\mathbf{r}} \mapsto \frac{\mathbf{r}_1 + g - (\mathbf{r}_2 + g)}{\|\mathbf{r}_1 + g - (\mathbf{r}_2 + g)\|} = \hat{\mathbf{r}}, \tag{3.23}$$

we see that

$$H(\mathbf{r}_1 + g, \mathbf{r}_2 + g; \mathbf{p}_1, \mathbf{p}_2) = H(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2), \tag{3.24}$$

so that $H$ is translation invariant.

The moment map associated to the group action is $(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2) \mapsto \mathbf{p}_1 + \mathbf{p}_2$, as in the N2BP. Picking a constant value $\alpha$ of $\Phi_0$, the reduced space becomes $M_1 := \Phi_0^{-1}(\alpha)/(\mathbb{R}^3, +) \cong \mathbb{R}^6$, where we have chosen the same coordinates $(\mathbf{r}, \mathbf{p})$ as in (2.77). Using the relations (2.78), we may find the expression for the new Hamiltonian $H_1$ on the reduced phase space $M_1$. We have

$$
\begin{aligned}
H_1(\mathbf{r}, \mathbf{p}; \alpha) &= H\left(\frac{1}{2}\mathbf{r}, -\frac{1}{2}\mathbf{r}; \frac{m_1}{M}\,\alpha + \mathbf{p}, \frac{m_2}{M}\,\alpha - \mathbf{p}\right) \\
&= \frac{1}{2}\left[\frac{\|\mathbf{p}\|^2}{\mu} + \frac{\|\alpha\|^2}{M}\right] \\
&\quad - \frac{1}{8c^2 {m_1}^3}\left\|\frac{m_1}{M}\,\alpha + \mathbf{p}\right\|^4 - \frac{1}{8c^2 {m_2}^3}\left\|\frac{m_2}{M}\,\alpha - \mathbf{p}\right\|^4 + \frac{e_1 e_2}{\|\mathbf{r}\|} \\
&\quad - \frac{e_1 e_2}{2c^2 m_1 m_2}\,\frac{1}{\|\mathbf{r}\|}\,\left[\left(\frac{m_1}{M}\,\alpha + \mathbf{p}\right)\cdot\left(\frac{m_2}{M}\,\alpha - \mathbf{p}\right)\right. \\
&\quad \left. + \left\{\left(\frac{m_1}{M}\,\alpha + \mathbf{p}\right)\cdot\hat{\mathbf{r}}\right\}\left\{\left(\frac{m_2}{M}\,\alpha - \mathbf{p}\right)\cdot\hat{\mathbf{r}}\right\}\right]. \qquad (3.25)
\end{aligned}
$$

At this point, we encounter the first glaring difference between the D2BP and the N2BP in terms of symplectic reduction. In the N2BP we found that keeping track of the total momentum $\alpha$, even when nonzero, had no physical consequences. The total momentum was only realized as an additive constant in the reduced Hamiltonian; when taking derivatives in order to obtain the equations of motion, this constant term would not contribute. In the D2BP, by contrast, we find terms containing $\alpha \cdot \hat{\mathbf{r}}/\|\mathbf{r}\|$ and $\alpha \cdot \mathbf{p}/\|\mathbf{r}\|$. These terms will have a significant impact on the equations of motion, so that the equations in the case $\alpha \neq \mathbf{0}$ will differ significantly from those with $\alpha = \mathbf{0}$. The physical reason for such a difference is not altogether clear, since the motion should be the same regardless of the overall translational motion of the system.

In the center-of-mass frame, the total momentum $\mathbf{p}_1 + \mathbf{p}_2$ is zero, hence $\alpha \equiv 0$, and so the above reduces to

$$H_1(\mathbf{r}, \mathbf{p}; \mathbf{0}) = \frac{\|\mathbf{p}\|^2}{2\mu} - \frac{\|\mathbf{p}\|^4}{8c^2 m^3} + \frac{e_1 e_2}{\|\mathbf{r}\|} + \frac{e_1 e_2}{2c^2 \mu M} \frac{1}{\|\mathbf{r}\|} \left[ \|\mathbf{p}\|^2 + (\mathbf{p} \cdot \hat{\mathbf{r}})^2 \right], \quad (3.26)$$

where

$$\frac{1}{m^3} := \frac{1}{{m_1}^3} + \frac{1}{{m_2}^3} \quad (3.27)$$

defines a "pseudo-reduced mass" $m$.

### 3.1.2 Rotation Invariance

As in the N2BP, we focus on the diagonal action of $SO(3)$, given by $S_g(\mathbf{r}, \mathbf{p}) = (g\mathbf{r}, g\mathbf{p})$ for $g \in SO(3)$. To determine if the Darwin system is rotationally invariant, we must compute $H_1 \circ S_g$, where $H_1$ is given by (3.25). Immediately we encounter our second obstacle: if we naively write $H_1 \circ S_g$ as $H_1(g\mathbf{r}, g\mathbf{p}; \alpha)$, then terms such as $\|(m_1/M)\alpha - \mathbf{p}\|^4$ will *not* be invariant. From our discussion of the N2BP in the previous chapter, it would appear that $g$ is applied only to $\mathbf{p}$ in such terms, and not to $\alpha$. If this is the case, $g$ will not factor out of the product $((m_1/M)\alpha - g\mathbf{p}) \cdot ((m_1/M)\alpha - g\mathbf{p})$, and such terms will clearly not be invariant.

On physical grounds there is no reason to believe the Darwin system fails to have rotational invariance. Rotational invariance is clearly in evidence in the case $\alpha \equiv \mathbf{0}$: using the fact that, $g^T g = \mathrm{id}$ for $g \in SO(3)$, and that

$\|\mathbf{p}\|^2 = \mathbf{p}^T\mathbf{p}$ and $\|\mathbf{r}\| = \sqrt{\mathbf{r}^T\mathbf{r}}$, we find

$$
\begin{aligned}
H_1(g\mathbf{r}, g\mathbf{p}; \alpha \equiv 0) &= \frac{1}{2\mu}\mathbf{p}^T g^T g\mathbf{p} - (\ldots)\left(\mathbf{p}^T g^T g\mathbf{p}\right)^2 + \frac{e_1 e_2}{\sqrt{\mathbf{r}^T g^T g\mathbf{r}}} \\
&\quad -(\ldots)\frac{1}{\sqrt{\mathbf{r}^T g^T g\mathbf{r}}}\left[\mathbf{p}^T g^T g\mathbf{p} + \left(\mathbf{p}^T g^T g\hat{\mathbf{r}}\right)^2\right] \\
&= H_1(\mathbf{r}, \mathbf{p}; \alpha \equiv 0),
\end{aligned} \tag{3.28}
$$

so that $H_1(\mathbf{r}, \mathbf{p}; \alpha \equiv 0)$ is invariant under the diagonal action of $SO(3)$. Thus the Darwin system is clearly invariant in the center-of-mass frame. There is no reason, however, that it should not also be rotationally invariant in a frame with nonzero total linear momentum. It seems that the method of symplectic reduction *forces* the choice of the center-of-mass frame in order to exploit rotational invariance, when there is no physical reason to do so.

A possible resolution to this seemingly unnatural state of affairs lies in a feature which was glossed over in the treatment of symplectic reduction of the N2BP.[4] Specifically, in the moment map construction for $(\mathbb{R}^3, +)$, the momentum $\mathbf{p}$ is reinterpreted as an element of the dual of the Lie algebra of $(\mathbb{R}^3, +)$, *cf.* expression (2.64). As shown in (2.65), $\alpha$ is an element of the same space. Therefore, as $g$ acts on $\mathbf{p}$, it must also act on $\alpha$. If we express this action as a matrix $(g)$ operating on a vector $\mathbf{p} \in \mathbb{R}^3$, then we must also express $\alpha$ as an element of $\mathbb{R}^3$, and $(g)$ must operate on it in the same fashion. The "constant" $\alpha$ cannot be left idle as we tacitly assumed in the N2BP, but must

[4]This is an omission not only of the treatment given here, but of several other treatments which have served as the source for the discussion presented in the previous chapter.

be interpreted in its proper space and is therefore subject to all the group actions upon that space. We could perhaps write this more clearly by tracing everything back to the orginal Hamiltonian $H$:

$$\begin{aligned} H_1 \circ S_g(\mathbf{r}, \mathbf{p}; \alpha) &= H \circ \bar{S}_g \left( \frac{1}{2}\mathbf{r}, -\frac{1}{2}\mathbf{r}; \frac{m_1}{M}\, \alpha + \mathbf{p}, \frac{m_2}{M}\, \alpha - \mathbf{p} \right) \\ &= H \left( \frac{1}{2} g\mathbf{r}, -\frac{1}{2} g\mathbf{r}; g \left( \frac{m_1}{M}\, \alpha + \mathbf{p} \right), g \left( \frac{m_2}{M}\, \alpha - \mathbf{p} \right) \right), \end{aligned} \tag{3.29}$$

where $\bar{S}_g$ is a natural extension of $S_g$ to the space $\{(\mathbf{r}_1, \mathbf{r}_2; \mathbf{p}_1, \mathbf{p}_2)\}$:

$$\begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \end{pmatrix} \mapsto \left( \begin{array}{c|c|c|c} g & & & \\ \hline & g & & \\ \hline & & g & \\ \hline & & & g \end{array} \right) \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \end{pmatrix} . \tag{3.30}$$

In this form it is a straightforward calculation to show that $H_1 \circ S_g = H_1$, and $H_1$ *is* rotationally invariant even in the case $\alpha \neq \mathbf{0}$.

The moment map $\Phi_1$ associated to the diagonal $SO(3)$ action is $(\mathbf{r}, \mathbf{p}) \mapsto (\mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T) \cong \mathbf{p} \times \mathbf{r}$, as in the N2BP (*cf.* (2.87)). Picking a constant value $\beta$ of $\Phi_1$, the reduced space becomes $M_2 := \Phi_1^{-1}(\beta)/SO(3)$, where we choose the same coordinates $(r, p_r)$ as in (2.106):

$$r := \|\mathbf{r}\|, \quad p_r := \frac{\mathbf{p} \cdot \mathbf{r}}{\|\mathbf{r}\|}, \tag{3.31}$$

so that $M_2 \cong \mathbb{R}^{>0} \times \mathbb{R}$.

Here we encounter another subtlety of symplectic reduction. As is clear from the discussion following (2.106), the coordinates above are obtained from $(\mathbf{r}, \mathbf{p})$ by application of a rotation $g_0 \in SO(3)$. In the discussion of the

N2BP, this rotation was naively applied to $(\mathbf{r}, \mathbf{p})$ alone; but if the reasoning above is correct, then $g_0$ should also be applied to $\alpha$. Thus in the switch to the new coordinates, $\alpha$ will transform to a new vector $\bar{\alpha} := g_0\alpha$. This seems plausible on physical grounds: $\alpha$ is the value of the total linear momentum, and a rotation of the system should entail a rotation of the total momentum. If, by contrast, we argue on purely mathematical grounds, then $\alpha$ is merely a parameter with three components, akin to other constants in the Hamiltonian. From that point of view, there is no *a priori* reason it should be subject to the rotation $g_0$. But if this is the case, then the Hamiltonian will *not* be invariant under the rotation $g_0$, unless $\alpha \equiv \mathbf{0}$.

We are in the fortuitous position of having learned something either way. If the Hamiltonian is to remain invariant for arbitrary values $\alpha$ of the moment map for the first reduction, then this parameter must be subject to the action of the group under which the second reduction takes place. Such requirements are not at all clear from the preponderance of literature on the topic of symplectic reduction. On the other hand, if one decides that there is no mathematically justifiable reason to subject a free parameter to the action of a group on the phase space coordinates (a Hamiltonian is a Hamiltonian is a Hamiltonian — we could choose to forget that any reduction has been performed to arrive at the functional form we now have), then we find that only *one* value of the preceding moment map, namely $\alpha \equiv \mathbf{0}$, allows for further reduction. In this case, the mathematics *does* choose the center-of-mass frame, with zero total momentum. The question of which route to choose has no clear answer, and begs the question of how well the mathematics of symplectic reduction is intended to mimic the methods and intuitions of physics.

The origin of such distinctions between the symmetries of the N2BP and the D2BP must certainly lie in the differences between Galilean and Lorentz invariance. The N2BP is a Galilean invariant theory, while the full relativistic system is Lorentz invariant. The D2BP, being not quite the N2BP, but neither quite fully relativistic, occupies a sort of nether region between the two. Further study of the symmetries of the Darwin system should study to what degree, if at all, the system retains any of the original Lorentz symmetry, particularly in orders of $v/c$. The present study, however, is concerned solely with the procedure of symplectic reduction, and so we will set aside further discussion of this issue.

For simplicity we may concentrate on the case $\alpha \equiv \mathbf{0}$. We obtain the following Hamiltonian:

$$\begin{aligned}
H_2(r, p_r; \alpha \equiv 0) &= H_1\left(\begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} p_r \\ -\hat{\beta}/r \\ 0 \end{pmatrix}; \alpha \equiv 0\right) \\
&= \frac{1}{2\mu}\left(p_r^2 + \frac{\hat{\beta}^2}{r^2}\right) - \frac{1}{8c^2m^3}\left(p_r^2 + \frac{\hat{\beta}^2}{r^2}\right)^2 \\
&\quad + \frac{e_1e_2}{r} - \frac{e_1e_2}{2c^2\mu M}\,\frac{1}{r}\left[2p_r^2 + \frac{\hat{\beta}^2}{r^2}\right],
\end{aligned} \tag{3.32}$$

where I've omitted the calculation — it amounts to merely expanding the dot products containing momenta in terms of the new expressions $(p_r, -\hat{\beta}/r, 0)$

and replacing $\|\mathbf{r}\|$ by $r$. From now on, we shall treat only the case $\alpha \equiv \mathbf{0}$. This condition will be assumed without any specific mention in what follows.

## 3.2 Dimensionless Equations

In order to analyze the equations of motion, it helps to first write the system in dimensionless form. This will allow us to be concrete in the evaluation of relative magnitudes of parameters governing the solutions. In our Hamiltonian (3.32), we have a characteristic velocity $c$, as well as a characteristic mass $\mu$. We may thus measure $H_2$ in units of $\mu c^2$. So we will work with a dimensionless Hamiltonian $H$ given by

$$\begin{aligned} \tilde{H} \; &:= \; \frac{1}{\mu c^2} H_2 = \frac{1}{2}\left(\left(\frac{p_r}{\mu c}\right)^2 + \left(\frac{\hat{\beta}}{\mu c r}\right)^2\right) - \frac{1}{8}\,\frac{\mu^3}{m^3}\left(\left(\frac{p_r}{\mu c}\right)^2 + \left(\frac{\hat{\beta}}{\mu c r}\right)^2\right)^2 \\ &\quad + \frac{e_1 e_2}{\mu c^2 r} - \frac{e_1 e_2}{2 M c^2}\,\frac{1}{r}\left[2\left(\frac{p_r}{\mu c}\right)^2 + \left(\frac{\hat{\beta}}{\mu c r}\right)^2\right]. \end{aligned} \tag{3.33}$$

This motivates the following definitions:

$$\tilde{p}_r \; := \; \frac{p_r}{\mu c}\,, \qquad \tilde{r} \; := \; \frac{\mu c^2}{e_1 e_2}\, r\,, \qquad \text{and} \qquad \tilde{\beta} \; := \; \frac{c}{e_1 e_2}\,\hat{\beta}\,, \tag{3.34}$$

all of which are dimensionless. Writing the dimensionless Hamiltonian $\tilde{H}$ in terms of these quantities, we have

$$\begin{aligned} \tilde{H}(\tilde{r}, \tilde{p}_r) &= \frac{1}{2}\left(\tilde{p}_r^2 + \left(\frac{\tilde{\beta}}{\tilde{r}}\right)^2\right) - \frac{1}{8}\frac{\mu^3}{m^3}\left(\tilde{p}_r^2 + \left(\frac{\tilde{\beta}}{\tilde{r}}\right)^2\right)^2 \\ &\quad + \frac{1}{\tilde{r}} - \frac{1}{2}\,\frac{\mu}{M}\,\frac{1}{\tilde{r}}\left[2\tilde{p}_r^2 + \left(\frac{\tilde{\beta}}{\tilde{r}}\right)^2\right], \end{aligned} \tag{3.35}$$

so that the Hamiltonian is given in terms of the dimensionless variables $(\tilde{r}, \tilde{p}_r)$ and the dimensionless parameters $\mu/M$, $\mu^3/m^3$, and $\tilde{\beta}$.

### 3.2.1 Hamilton's Equations and Dimensionless Time

In order that the equations be dimensionless, we must establish a dimensionless time variable. Starting with Hamilton's equations,

$$\frac{dr}{dt} = \frac{\partial H}{\partial p_r}, \qquad \frac{dp_r}{dt} = -\frac{\partial H}{\partial r} \ , \tag{3.36}$$

we define a dimensionless time variable

$$\tau \ := \ \frac{t}{T} \ , \tag{3.37}$$

where $T$ has the units of time. Then

$$\frac{d}{dt} = \frac{d\tau}{dt}\frac{d}{d\tau} = \frac{1}{T}\frac{d}{d\tau} \ . \tag{3.38}$$

If we substitute these relations and the definitions (3.34) into Hamilton's equations (3.36), noting $\tilde{H} = H/\mu c^2$, we have

$$\begin{aligned} \frac{1}{T}\frac{e_1 e_2}{\mu c^3}\frac{d\tilde{r}}{d\tau} &= \frac{\partial \tilde{H}}{\partial \tilde{p}_r} \\ \frac{1}{T}\frac{e_1 e_2}{\mu c^3}\frac{d\tilde{p}_r}{d\tau} &= -\frac{\partial \tilde{H}}{\partial \tilde{r}} \ . \end{aligned} \tag{3.39}$$

Choosing

$$T \ := \ \frac{e_1 e_2}{\mu c^2} \ , \tag{3.40}$$

$\tau$ is dimensionless as expected, and so we may now work with entirely dimensionless equations.

### 3.2.2 Equations of Motion

We may now use the dimensionless equations of Hamilton and $\tilde{H}$ to obtain the equations of motion in dimensionless form. We have

$$
\begin{aligned}
\frac{d\tilde{r}}{d\tau} &= \frac{\partial \tilde{H}}{\partial \tilde{p}_r} = \tilde{p}_r \left[ 1 - \frac{1}{2}\frac{\mu^3}{m^3}\left(\tilde{p}_r^2 + \frac{\tilde{\beta}^2}{\tilde{r}^2}\right) - 2\frac{\mu}{M}\frac{1}{\tilde{r}} \right] \\
\frac{d\tilde{p}_r}{d\tau} &= -\frac{\partial \tilde{H}}{\partial \tilde{r}} = \frac{1}{\tilde{r}} \left[ \left(1 - \frac{\mu}{M}\tilde{p}_r^2\right) + \frac{1}{\tilde{r}}\left(\tilde{\beta}^2 - \frac{1}{2}\frac{\mu^3}{m^3}\tilde{p}_r^2\right) \right. \\
&\qquad \left. - \frac{1}{\tilde{r}^2}\frac{3}{2}\frac{\mu}{M}\tilde{\beta}^2 - \frac{1}{\tilde{r}^3}\frac{1}{2}\frac{\mu^3}{m^3}\tilde{\beta}^2 \right] .
\end{aligned}
\tag{3.41}
$$

**NB:** From now on, I will drop the tildes — all variables will henceforth be dimensionless unless stated otherwise. The notation $\dot{r}$ will refer to differentiation of the dimensionless $r$-variable with respect to the dimensionless time variable $\tau$, and likewise for other quantities.

## 3.3 Quadrature

Darwin in his original paper proceeded from the Euler-Lagrange equations to a solution of the same order of approximation as the system itself. In this section we take a different route, proceeding from the Hamiltonian to an implicit, but exact, quadrature of the system akin to that found in the N2BP. Starting with the dimensionless Hamiltonian (3.35), the condition

$H = \text{const} =: E$ implies

$$0 = \left(-\frac{\mu^3}{8m^3}\right) p_r^4 + \left(\frac{1}{2} - \frac{\mu^3}{4m^3}\frac{\beta^2}{r^2} - \frac{\mu}{Mr}\right) p_r^2$$
$$+ \left(\frac{\beta^2}{2r^2} + \frac{1}{r} - \frac{\mu\beta^2}{2Mr^3} - \frac{\mu^3\beta^4}{8m^3r^4} - E\right) . \tag{3.42}$$

This is a quadratic equation in $p_r^2$, and so application of the quadratic formula gives

$$p_r^2 = -y_1 \left[-y_2 \pm \sqrt{y_2^2 + y_3}\right] , \tag{3.43}$$

where the functions $y_i$ depend on $1/r$ and are given by

$$\begin{aligned} y_1 &:= \frac{\mu^3}{8m^3} , \\ y_2 &:= \frac{1}{2} - \frac{\mu^3}{4m^3}\frac{\beta^2}{r^2} - \frac{\mu}{Mr} , \\ y_3 &:= \frac{\mu^3}{2m^3}\left(\frac{\beta^2}{2r^2} + \frac{1}{r} - \frac{\mu\beta^2}{2Mr^3} - \frac{\mu^3\beta^4}{8m^3r^4} - E\right) . \end{aligned} \tag{3.44}$$

Reality of $p_r$ requires that

$$\left[-y_2 \pm \sqrt{y_2^2 + y_3}\right] < 0. \tag{3.45}$$

The term under the square root must of course remain positive, so that we require

$$|y_3| \geq -y_2^2 \tag{3.46}$$

at all times. There are two main cases, based on whether $y_2$ is positive or negative.

**Case 1**: $y_2 > 0$. When $y_2$ is positive (*i.e.* $y_2 = |y_2|$), then $[\ldots] < 0$ always obtains when the negative root $-y_2 - \sqrt{y_2^2 + y_3}$ is chosen, as long as the condition (3.46) is met. If the positive root $-y_2 + \sqrt{y_2^2 + y_3}$ is chosen, then the term will be negative when

$$\sqrt{y_2^2 + y_3} < |y_2| \ , \tag{3.47}$$

which implies

$$0 > y_3 > -y_2^2 \ . \tag{3.48}$$

Therefore *two* real roots are ensured when $y_2 > 0$ and $0 > y_3 > -y_2^2$.

**Case 2**: $y_2 < 0$. In this case, $-y_2 = |y_2|$, and so the condition $[\ldots] < 0$ prohibits the positive root. Only the negative root $-y_2 - \sqrt{y_2^2 + y_3}$ is allowed, which will be negative provided $y_3 > 0$.

Thus we have real $p_r$

$$\begin{aligned} &\text{for} \qquad y_2 > 0 \qquad \text{when} \qquad y_3 > 0 \qquad \text{or} \qquad 0 > y_3 > -y_2^2 \ , \\ &\text{and for} \qquad y_2 < 0 \qquad \text{when} \qquad y_3 > 0 \ . \end{aligned} \tag{3.49}$$

Taken together, we are ensured real values of $p_r$ when

$$y_3 > 0, \qquad \text{that is,} \qquad \frac{\beta^2}{2r^2} + \frac{1}{r} - \frac{\mu\beta^2}{2Mr^3} - \frac{\mu^3\beta^4}{8m^3r^4} > E \ . \tag{3.50}$$

Assuming this condition holds, and choosing the negative root common to both cases, then another application of the quadratic formula yields

$$p_r = \pm\sqrt{-y_1 \left[-y_2 - \sqrt{y_2^2 + y_3}\right]} \ . \tag{3.51}$$

Referring back to the equations of motion (3.41), substitution into the equation for $dr/d\tau$ yields

$$\frac{dr}{d\tau} = \pm\sqrt{-y_1\left[-y_2 - \sqrt{y_2^2 + y_3}\right]} \left\{1 - \frac{\mu^3}{2m^3}\left(-y_1\left[-y_2 - \sqrt{y_2^2 + y_3}\right] + \frac{\beta^2}{r^2}\right) - \frac{2\mu}{Mr}\right\} . \quad (3.52)$$

Integration yields

$$\tau - \tau_0 = \pm\int_{r_0}^{r} \mathrm{d}r' \quad \left[\frac{1}{\sqrt{-y_1\left[-y_2 - \sqrt{y_2^2 + y_3}\right]}} \cdot \frac{1}{\left\{1 - \frac{\mu^3}{2m^3}\left(-y_1\left[-y_2 - \sqrt{y_2^2 + y_3}\right] + \frac{\beta^2}{r'^2}\right) - \frac{2\mu}{Mr'}\right\}}\right] , \quad (3.53)$$

giving a quadrature of the D2BP. In the integrand, the $y_i$ are functions of $1/r'$. The RHS is a function of $r$, so that inversion yields $r(\tau)$.

## 3.4 Fixed Points of the Equations of Motion

We are now in a position to look for points where

$$\dot{r} = 0, \qquad \dot{p}_r = 0, \quad (3.54)$$

that is, the fixed points of the motion, corresponding to circular orbits. Looking at equation (3.41), the first of the two expressions suggests that taking

$p_r \equiv 0$ might be a promising route. In this case the equations reduce to

$$\dot{r} = 0$$

$$\dot{p}_r = \frac{1}{r^2}\left[1 + \frac{1}{r}\,\beta^2 - \frac{1}{r^2}\,\frac{3}{2}\frac{\mu}{M}\,\beta^2 - \frac{1}{r^3}\,\frac{1}{2}\frac{\mu^3}{m^3}\,\beta^2\right]. \tag{3.55}$$

In order for the second equation to give zero, since $1/r^2 \neq 0$, the term in brackets must be zero. Setting

$$R \;:=\; \frac{1}{r}\,, \tag{3.56}$$

we want to solve the equation

$$1 + R\,\beta^2 - R^2\,\frac{3}{2}\frac{\mu}{M}\,\beta^2 - R^3\,\frac{1}{2}\frac{\mu^3}{m^3}\,\beta^2 \overset{\heartsuit}{=} 0, \tag{3.57}$$

which is a cubic equation in $R$.

### 3.4.1 General Solution to the Cubic

Luckily, the cubic has an explicit general solution, albeit a fair bit more complicated than the quadratic. Suppose we wish to solve

$$a_3 z^3 + a_2 z^2 + a_1 z + a_0 = 0. \tag{3.58}$$

To start, define the quantities

$$\begin{aligned} q \;&:=\; \frac{1}{3}\,\frac{a_1}{a_3} - \frac{1}{9}\left(\frac{a_2}{a_3}\right)^2 \;; \\ r \;&:=\; \frac{1}{6}\left(\frac{a_1 a_2}{a_3^2} - 3\frac{a_0}{a_3}\right) - \frac{1}{27}\left(\frac{a_2}{a_3}\right)^3 \;. \end{aligned} \tag{3.59}$$

There are three possible scenarios:

- *CASE 1 (reducible):* $q^3 + r^2 > 0$ — in this situation, there is one real root and two complex roots.
- *CASE 2 (reducible):* $q^3 + r^2 = 0$ — all the roots are real, and at least two are equal.
- *CASE 3 (irreducible):* $q^3 + r^2 < 0$ — all the roots are real.

In the case of reducible roots (the first two scenarios), we further define the quantities

$$\begin{aligned} s_1 &:= \left(r + \sqrt{q^3 + r^2}\right)^{1/3} \; ; \\ s_1 &:= \left(r - \sqrt{q^3 + r^2}\right)^{1/3} \; . \end{aligned} \tag{3.60}$$

The roots $z_1, z_2, z_3$ of the cubic are then given by

$$\begin{aligned} z_1 &= s_1 + s_2 - \frac{1}{3}\frac{a_2}{a_3} \; ; \\ z_2 &= -\frac{1}{2}(s_1 + s_2) - \frac{1}{3}\frac{a_2}{a_3} + i\,\frac{\sqrt{3}}{2}(s_1 - s_2) \; ; \\ z_3 &= -\frac{1}{2}(s_1 + s_2) - \frac{1}{3}\frac{a_2}{a_3} - i\,\frac{\sqrt{3}}{2}(s_1 - s_2) \; . \end{aligned} \tag{3.61}$$

In the case of irreducible roots (the third case above), we define

$$\theta \; := \; \arccos\left(\frac{-r}{\sqrt{-q^3}}\right) . \tag{3.62}$$

Then one obtains the roots from the relations

$$z_1 = -2\sqrt{-q}\ \cos\left(\frac{\theta}{3}\right) - \frac{1}{3}\frac{a_2}{a_3}\ ;$$

$$z_2 = -2\sqrt{-q}\ \cos\left(\frac{\theta}{3} + \frac{2\pi}{3}\right) - \frac{1}{3}\frac{a_2}{a_3}\ ;$$

$$z_3 = -2\sqrt{-q}\ \cos\left(\frac{\theta}{3} - \frac{2\pi}{3}\right) - \frac{1}{3}\frac{a_2}{a_3}\ . \tag{3.63}$$

It ain't pretty, but it works. . . .

### 3.4.2 Fixed Points and the Cubic

In the equation before us, eqn (3.57), we have

$$a_0 = 1, \qquad a_1 = \beta^2, \qquad a_2 = -\frac{3}{2}\frac{\mu}{M}\beta^2, \qquad a_3 = -\frac{1}{2}\frac{\mu^3}{m^3}\beta^2. \tag{3.64}$$

This gives

$$q = -\frac{2}{3}\frac{m^3}{\mu^3}\left(1 + \frac{3}{2}\frac{m^3}{M^2\mu}\right)\ ;$$

$$r = -\frac{m^3}{\mu^3}\left(\frac{m^3}{\mu^2 M} - \frac{1}{\beta^2} + \frac{m^6}{\mu^3 M^3}\right)\ . \tag{3.65}$$

Cubing and squaring the respective terms, then adding, yields

$$q^3 + r^2 = \frac{m^6}{\mu^{10}M^3}\frac{1}{\beta^4}\Bigg[\ \mu^4 M^3 - \left(\frac{8}{27}m^3\mu M^3\beta^4 + \tfrac{1}{3}m^6\beta^4 + 2m^3\mu^2M^2\beta^2 + 2m^6\mu\beta^2\right)\Bigg] . \tag{3.66}$$

Recall from the preceding section that the behavior of the quantity $q^3 + r^2$ determines the nature of the roots of the cubic. We are certainly guaranteed *at least* one *real* solution. In that sense, we were done before we started to calculate $q^3 + r^2$. However, it is not clear at the moment how to interpret complex solutions for $R := 1/r$. Thus we should explore the conditions that ensure *all real* roots. In that case, we want

$$q^3 + r^2 \overset{\heartsuit}{\leq} 0. \tag{3.67}$$

Since all the mass terms are positive, and $\beta$ enters into $q^3 + r^2$ only in combinations of $\beta^2 \geq 0$, then to satisfy the above condition it is enough to require[5]

$$\left[\mu^4 M^3 - \beta^2 \left(\frac{8}{27} m^3 \mu M^3 \beta^2 + \tfrac{1}{3} m^6 \beta^2 + 2m^3 \mu^2 M^2 + 2m^6 \mu\right)\right] \overset{\heartsuit}{\leq} 0. \tag{3.68}$$

If $\beta$ were allowed to be arbitrary, we would be finished. We would always be able to choose $\beta$ large enough to ensure that the quantity $[-2\beta^2 \cdot$ (positive terms)$]$ be larger than $\mu^4 M^3$, and hence the entire expression could be made negative. But for the Darwin Hamiltonian to be a valid approximation to a fully relativistic system, we require that $(v/c)$ be in some sense "small". Just how small is a matter of taste, depending on how well we want the Darwin system to mimic a relativistic one, but certainly we cannot allow $(v/c)$ to be arbitrarily large. This then will restrict our allowable range for $\beta$.

To see what restrictions might be placed on $\beta$, let us write the inequality (3.68) in terms of the original masses of the particles. For the individual

[5]At first glance, it might seem odd that the condition for fixed points of the equations of motion for an electromagnetic system does not depend on the charge. Recall, however, that $\beta$ (really $\tilde{\beta}$) actually does depend on the charge: $\tilde{\beta} := (c/e_1 e_2)\,\hat{\beta}$. What is interesting is that the charge only enters into the condition in combination with the angular momentum, not separately.

terms we have

$$
\begin{aligned}
\mu^4 M^3 &= \tfrac{m_1{}^4 m_2{}^4}{m_1+m_2}\ ; & \qquad m^3\mu M^3 &= m_1{}^4 m_2{}^4 \cdot \tfrac{(m_1+m_2)^2}{m_1{}^3+m_2{}^3}\ ; \\
m^6 &= \tfrac{m_1{}^6 m_2{}^6}{(m_1{}^3+m_2{}^3)^2}\ ; & \qquad m^3\mu^2 M^2 &= \tfrac{m_1{}^5 m_2{}^5}{m_1{}^3+m_2{}^3}\ ;
\end{aligned}
\tag{3.69}
$$

and

$$
m^6\mu = \frac{m_1{}^7 m_2{}^7}{(m_1{}^3+m_2{}^3)^2(m_1+m_2)}\ . \tag{3.70}
$$

Then (3.68) becomes

$$
\begin{aligned}
0 \ \overset{\heartsuit}{\geq}\ & \Bigg[(m_1{}^3+m_2{}^3)^2 - \beta^2 \Bigg(\frac{8}{27}\beta^2(m_1{}^3+m_2{}^3)(m_1+m_2)^3 \\
& \tfrac{1}{3}\beta^2 m_1{}^2 m_2{}^2(m_1+m_2) + 2m_1 m_2(m_1{}^3+m_2{}^3)^2(m_1+m_2) \\
& +\ \ 2m_1{}^3 m_2{}^3\Bigg)\Bigg]\ ,
\end{aligned}
\tag{3.71}
$$

where I have omitted a factor

$$
\frac{m_1{}^4 m_2{}^4}{(m_1{}^3+m_2{}^3)^2(m_1+m_2)} \tag{3.72}
$$

multiplying the entire right-hand side. This expression may be factored in the following manner:

$$
\begin{aligned}
\left(1-\frac{8}{27}\beta^4\right)m_1{}^6 + \left(2-\frac{16}{27}\beta^4\right)m_1{}^3 m_2{}^3 + \left(1-\frac{8}{27}\beta^4\right)m_2{}^6 - \ldots \ \overset{\heartsuit}{\leq}\ 0 \\
\left(1-\frac{8}{27}\beta^4\right)\left(m_1{}^3+m_2{}^3\right)^2 - (\text{positive terms})\ \overset{\heartsuit}{\leq}\ 0.
\end{aligned}
\tag{3.73}
$$

This will certainly hold if we require

$$
1-\frac{8}{27}\beta^4 \overset{\heartsuit}{\leq} 0, \qquad \text{that is,} \qquad \sqrt[4]{\frac{27}{8}} \overset{\heartsuit}{\leq} \beta. \tag{3.74}
$$

Writing this in terms of the original angular momentum $\hat{\beta}$, this becomes

$$\sqrt[4]{\frac{27}{8}} \overset{\heartsuit}{\leq} \frac{c}{e_1 e_2}\hat{\beta}. \tag{3.75}$$

By the conservation of angular momentum, *i.e.* $SO(3)$ invariance, we may use the relation

$$\hat{\beta} = \mu r^2 \dot{\phi}, \tag{3.76}$$

where $\dot{\phi}$ is the rate of rotation. Then the condition we want is

$$\sqrt[4]{\frac{27}{8}} \overset{\heartsuit}{\leq} \frac{c}{e_1 e_2}\mu r^2 \dot{\phi}. \tag{3.77}$$

Thus if the above condition is satisfied, either in the form (3.75) or the form (3.77), we are ensured three *real* roots for the radius of the fixed points. The fixed points all have the form $(r, p_r) = (r_i, 0)$, for $r_i$ the three solutions to the cubic in $R := 1/r$.

### 3.4.3 Physicality of the Root Conditions and an Estimate for $v/c$

Let's examine just how physical the bounds arrived at in the last section really are. To be concrete, suppose we look at a proton and electron separated by the Bohr radius. Substituting the Bohr radius for $r$ and computing the reduced mass of the proton-electron system, the condition (3.77) becomes[6]

$$\sqrt[4]{\frac{27}{8}} \overset{\heartsuit}{\leq} \frac{3\times 10^{10}\,\text{cm/sec}}{(1.6\times 10^{-19}\,\text{coul})^2}\cdot\frac{(1\text{coul})^2}{(3\times 10^{9}\text{esu})^2}$$

$$\frac{(1.67\times 10^{-24}\,\text{g})(9.11\times 10^{-28}\,\text{g})}{(1.67\times 10^{-24}\,\text{g})+(9.11\times 10^{-28}\,\text{g})}\cdot \quad (5.292\times 10^{-9}\,\text{cm})^2\,\dot{\phi}$$

$$\sim\ 10^{-16}[\text{sec}]\cdot\dot{\phi}\,. \tag{3.78}$$

Further estimating $27/8 \sim 3$, so that $\sqrt{3}\sim 1$ and hence $\sqrt[4]{3}\sim 1$, we then have the condition

$$10^{16} \overset{\heartsuit}{\lesssim} \dot{\phi}\ [\text{sec}]. \tag{3.79}$$

To see how this translates into a condition on the ratio $(v/c)$, we first note that for the velocity of the reduced particle,

$$v^2 = \dot{r}^2 + r^2\dot{\phi}^2 = r^2\dot{\phi}^2\ , \tag{3.80}$$

where I have used the fact that at the fixed point $\dot{r}=0$. Hence

$$\dot{\phi} = \frac{v}{r}\ . \tag{3.81}$$

Using this in our relation (3.79), and again using the Bohr radius for $r$, we find

$$10^{16} \overset{\heartsuit}{\lesssim} \frac{v}{10^{-9}\,\text{cm}}\ [\text{sec}]\ , \qquad \text{or, rather,} \qquad 10^{25}\,\frac{\text{cm}}{\text{sec}} \overset{\heartsuit}{\lesssim} v. \tag{3.82}$$

In terms of the ratio $(v/c)$, and writing $c\sim 10^{10}$ (cm/sec), we obtain

$$10^{15} \overset{\heartsuit}{\lesssim} \frac{v}{c}\ . \tag{3.83}$$

[6]Note that we are using CGS units, so that charge is related to the other fundamental units by $Q^2 = ML^3/T^2$, and hence $[\text{esu}^2] = [\text{g}\cdot\text{cm}^3/\text{sec}^2]$.

This is obviously prohibitively large, implying that at a distance of roughly the Bohr radius, only one fixed point, *i.e.* one circular orbit, is possible. We might also investigate whether there are any distances other than the Bohr radius which might permit three fixed points. If we use instead the Debye radius

$$r_D = 6.9 \left(\frac{T}{n_0}\right)^{1/2} \text{ [cm]} \; , \tag{3.84}$$

where $T$ is given in Kelvin and $n_0$ is in cm$^{-3}$, then we may repeat the above calculation to arrive at an estimate which depends on density and temperature. We have

$$10^{16} \overset{\heartsuit}{\lesssim} \left(\frac{10^{-9}}{r_D}\right)^2 \dot{\phi} \text{ [sec]}, \qquad \text{so that} \qquad 10^{34} \overset{\heartsuit}{\lesssim} \left(\frac{T}{n_0}\right) \dot{\phi} \text{ [sec]}. \tag{3.85}$$

In terms of $(v/c)$,

$$10^{24} \overset{\heartsuit}{\lesssim} \left(\frac{T}{n_0}\right)^{1/2} \frac{v}{c} \; . \tag{3.86}$$

With this expression, we can actually get a rough idea of what kind of physical regimes would allow for three real roots. If we stipulate a level of accuracy of the Darwin approximation, *i.e.* stipulate a range for $(v/c)$, then the above expression puts bounds on the temperature and density regimes which permit such solutions. Assuming a value $(v/c) \sim 10^{-2}$, this would require $T/n_0$ to be greater than $(10^{26})^2 = 10^{52}$. Typical ranges of $T$ for plasmas are roughly $10^3$K (for the Earth's ionosphere) to $10^8$K (for fusion reactors); and typical densities are anywhere from 1cm$^{-3}$ (for interstellar gas) to $10^{14}$cm$^{-3}$ (for fusion reactors). Even the highest value of $T/n_0$ falls far short of the requirement above, thereby ensuring only the possibility of one circular orbit.

## 3.5 Comparison with the Kepler Problem

Recall for a moment some facts about the Kepler problem (*i.e.* the usual 2BP). Once one has reduced by all the symmetries, the Hamiltonian becomes

$$H = \frac{p^2}{2m} - \frac{\text{const}}{r} + \frac{\text{ang. mom.}}{\text{mass}\, r^2} \ . \tag{3.87}$$

In this case Hamilton's equations become, roughly,

$$\dot{r} = \frac{p}{m} \ , \qquad \dot{p} = \frac{\text{const}}{r^2}\left(\frac{\text{ang. mom.}}{\text{mass}} - 1\right) \ . \tag{3.88}$$

Searching for the fixed point of the equations of motion, then, amounts to setting the momentum $p$ equal to zero and finding the roots of a polynomial in $1/r$. In this case one simply gets one solution:

$$r_0 = \frac{\text{mass}}{\text{ang. mom.}} \ . \tag{3.89}$$

This corresponds to a *circular* orbit, *i.e.* an orbit at fixed radius. This solution is unique, given the values of the parameters of the system. Thinking of the Hamiltonian as

$$H = T(p) + V_{eff}(r), \tag{3.90}$$

then looking for the fixed points amounts to looking for the minimum of $V_{eff}$, since

$$\dot{p} = -\frac{\partial V_{eff}}{\partial r} \ . \tag{3.91}$$

In the D2BP, we follow essentially the same procedure and obtain conditions under which we are assured circular orbits. The difference, however, is that we find, in principle, the possibility of *three* radii for circular orbits. Why this should happen is not entirely clear. It is possible that it is merely an artifact of truncating the relativistic system to obtain a polynomial expression in the momenta, so that all resulting equations will also be polynomial and therefore have a finite number of roots. The fact that the conditions under which all three radii could exist are too restrictive for likely physical situations lends support to this idea.

# Chapter 4

# General Features of the Darwin System

## 4.1 Darwin Potentials

The different forms of the electromagnetic potentials in different gauges are crucial for the ensuing studies of the Darwin system. This section outlines the procedure for obtaining explicit expressions for $\phi$ and $\mathbf{A}$. We may summarize the procedure as follows: solve the electromagnetic system in a given gauge to obtain solutions in terms of retarded quantities; then find the instantaneous approximations to this system.[1] One simple approach is to start in Lorenz gauge, in which the equations for $\mathbf{A}$ and $\phi$ decouple, thereby allowing for explicit solutions in terms of the sources. Then it is a quick calculation to change from Lorenz gauge to Coulomb gauge, so that both the equations and solutions are transformed. In Coulomb gauge it is clear that $\phi$ solves Poisson's equation, and therefore mimics the electrostatic (*i.e.* "instantaneous") case. The transformation will also alter the equation for, and therefore form of, $\mathbf{A}$. The expression for $\mathbf{A}$ is still given in terms of retarded quantities. Approximating this by the corresponding instantaneous quantity yields the same

[1]Fortunately, Jackson [28] has outlined just such a procedure, by which one can change from gauge to gauge to obtain different expressions for the scalar and vector potentials as functions of the retarded times, and then take the instantaneous limit of these.

microscopic system which Darwin considered in his original paper [9], and so we arrive at the Darwin system which is the object of our study.[2] The method used here is essentially the same as that followed in Section 3.1, but reversing the sequence of changing gauge and truncating at second order.

More specifically, we may describe the procedure as follows. Start with **E** and **B** in terms of the potentials:

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \qquad \text{and} \qquad \mathbf{B} = \nabla \times \mathbf{A}. \tag{4.1}$$

Assuming Lorenz gauge, these satisfy decoupled wave equations:

$$\begin{aligned} \nabla^2\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} &= -4\pi\rho; \\ \nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} &= -\frac{4\pi}{c}\mathbf{j}. \end{aligned} \tag{4.2}$$

The solutions are well-known, and so will be omitted here. We then transform from Lorenz gauge to Coulomb gauge. The equations become

$$\begin{aligned} \nabla^2\phi &= -4\pi\rho; \\ \nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} &= -\frac{4\pi}{c}\left(\mathbf{j} - \frac{1}{4\pi}\nabla\frac{\partial\phi}{\partial t}\right). \end{aligned} \tag{4.3}$$

One may of course recognize that the equations of motion for the potentials are just the general equations

$$\begin{aligned} \nabla^2\phi + \frac{1}{c}\frac{\partial}{\partial t}(\nabla\cdot\mathbf{A}) &= -4\pi\rho, \\ \nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} - \nabla\left(\nabla\cdot\mathbf{A} - \frac{1}{c}\frac{\partial\phi}{\partial t}\right) &= -\frac{4\pi}{c}\mathbf{j}, \end{aligned} \tag{4.4}$$

[2] Jackson, in the same paper, points out in §VI.B that the essentially the same system was studied earlier by Heaviside [24], but Darwin seems not to have been aware of his work.

with $\nabla \cdot \mathbf{A} = 0$, the usual condition for Coulomb gauge. The solutions transform as well. The scalar potential tranforms into the well-known solution of Poisson's equation,

$$\phi_C(\mathbf{x}, t) = \int \mathrm{d}^3\mathbf{x}' \frac{\rho(\mathbf{x}', t)}{\|\mathbf{x} - \mathbf{x}'\|} , \tag{4.5}$$

and the vector potential takes a less-familiar form:

$$\begin{aligned} \mathbf{A}_C(\mathbf{x}, t) = \frac{1}{c} \int & \frac{\mathrm{d}^3\mathbf{x}'}{\|\mathbf{x} - \mathbf{x}'\|} \left[\mathbf{j}(\mathbf{x}', t') - \hat{\mathbf{r}}\hat{\mathbf{r}} \cdot \mathbf{j}(\mathbf{x}', t')\right]_{\text{ret}} \\ & + c \int \frac{\mathrm{d}^3\mathbf{x}'}{\|\mathbf{x} - \mathbf{x}'\|^3} \int_0^{\|\mathbf{x}-\mathbf{x}'\|/c} \mathrm{d}\tau \; \tau \left[3\hat{\mathbf{r}}\hat{\mathbf{r}} \cdot \mathbf{j}(\mathbf{x}', t - \tau) - \mathbf{j}(\mathbf{x}', t - \tau)\right], \end{aligned} \tag{4.6}$$

where $C$ stands for Coulomb, and "ret" denotes that the quantities in brackets are to be evaluated at the retarded time

$$t' = t - \frac{\|\mathbf{x} - \mathbf{x}'\|}{c} . \tag{4.7}$$

Jackson gives an equivalent form of the vector potential,

$$\mathbf{A}_C(\mathbf{x}, t) = -c\nabla \times \int \mathrm{d}^3\mathbf{x}' \int_0^{\|\mathbf{x}-\mathbf{x}'\|/c} \mathrm{d}\tau \; \tau \left( \frac{\hat{\mathbf{r}} \times \mathbf{j}(\mathbf{x}', t - \tau)}{\|\mathbf{x} - \mathbf{x}'\|^2} \right) , \tag{4.8}$$

which makes it clear that $\mathbf{A}$ does in fact have zero divergence.

We may now use either expression for $\mathbf{A}_C$ to obtain the instantaneous vector potential in Coulomb gauge. Using our expressions in terms of retarded times — it is simplest to use the form (4.6), though one could also use (4.8) — and letting $\mathbf{j}(\mathbf{x}', t-\tau) \to \mathbf{j}(\mathbf{x}', t)$, then direct calculation yields the so-called "quasistatic" vector potential in Coulomb gauge:

$$\mathbf{A}_C^{qs}(\mathbf{x},t) = \frac{1}{2c}\int \mathrm{d}^3\mathbf{x}' \frac{\mathbf{j}(\mathbf{x}',t) + \hat{\mathbf{r}}(\hat{\mathbf{r}}\cdot\mathbf{j}(\mathbf{x}',t))}{\|\mathbf{x}-\mathbf{x}'\|}\ . \tag{4.9}$$

Expressing the current in terms of delta functions yields the vector potential treated in the microscopic Darwin problem.

## 4.2 Darwin Field Equations

The Darwin system may be approached from a different direction. Instead of solving for the potentials in Lorenz gauge, then transforming to Coulomb gauge and taking the instantaneous limit, we may start with the field equations themselves, take the instantaneous limit of these, and then solve directly for the potentials.

Before proceeding, we first note that any vector may be written as the sum of divergence-free and curl-free vectors. In particular, we may so decompose the current: $\mathbf{j} = \mathbf{j}_l + \mathbf{j}_t$, where $l$ and $t$ denote 'longitudinal' and 'transverse', respectively. Jackson gives explicit forms for this decomposition:

$$\begin{aligned}\mathbf{j}_l &= -\frac{1}{4\pi}\nabla\int \mathrm{d}^3\mathbf{x}'\frac{\nabla'\cdot\mathbf{j}(\mathbf{x}',t)}{\|\mathbf{x}-\mathbf{x}'\|} = \frac{1}{4\pi}\nabla\int \mathrm{d}^3\mathbf{x}'\frac{1}{\|\mathbf{x}-\mathbf{x}'\|}\frac{\partial\rho(\mathbf{x}',t)}{\partial t}\\ &= \frac{1}{4\pi}\nabla\frac{\partial\phi}{\partial t}\ ,\end{aligned} \tag{4.10}$$

where the continuity equation has been used in the second step, and

$$\mathbf{j}_t = \frac{1}{4\pi}\nabla\times\left(\nabla\times\int \mathrm{d}^3\mathbf{x}'\frac{\mathbf{j}(\mathbf{x}',t)}{\|\mathbf{x}-\mathbf{x}'\|}\right). \tag{4.11}$$

If we then start with Maxwell's field equations

$$\nabla \cdot \mathbf{E} = 4\pi\rho \qquad \text{and} \qquad \nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j};$$

$$\nabla \cdot \mathbf{B} = 0 \qquad \text{and} \qquad \nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = \mathbf{0}; \tag{4.12}$$

we may take the equation

$$\nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j} \tag{4.13}$$

and rewrite this as

$$\begin{aligned} \nabla \times \mathbf{B} &= \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\frac{\partial}{\partial t}\left(-\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right) \\ &= \frac{4\pi}{c}\mathbf{j}_t - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} \ . \end{aligned} \tag{4.14}$$

Then we see that

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{j}_t \qquad \Longleftrightarrow \qquad \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} \sim 0. \tag{4.15}$$

If we take the defining characteristic of the Darwin system to be that it is an expansion up to second order in $v/c$, then a heuristic argument lends credence to the idea that the above equation describes the same physics as the Darwin approximation. Specifically, as is easily seen from the expression for $\mathbf{A}$ in Lorenz gauge, $\mathbf{A}$ scales as $1/c$.[3] Then the term $(1/c^2)(\partial^2\mathbf{A}/\partial t^2)$ is of

[3]That is not to say we are working *in* Lorenz gauge — on the contrary, the Darwin system, however described, employs Coulomb gauge. The above argument is only using the *familiarity* of the Lorenz form of the vector potential to determine how $\mathbf{A}$ scales with $c$.

order $1/c^3$, and so should setting it to zero is consistent with the Darwin approximation.

On the other hand, many authors take the lack of retardation as the defining characteristic of the Darwin approximation (*e.g.* [37, 49]). We may rewrite the $\mathbf{B}$ equation above as

$$\frac{4\pi}{c}\mathbf{j}_t = \nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = -\nabla^2 \mathbf{A} \tag{4.16}$$

in Coulomb gauge. From the fact that the above is simply Poisson's equation (in contradistinction to the wave equations of Lorenz gauge, involving time-derivatives), it is clear that there are no retardation effects.

The solution to Poisson's equation is well-known, and so we may solve directly for $\mathbf{A}$ in terms of $\mathbf{j}$. Using our explicit expression for $\mathbf{j}_t$ in terms of $\mathbf{j}$, and applying the Green's function for Poisson's equation and the expression (4.11) for $\mathbf{j}_t$, we have

$$\begin{aligned}\mathbf{A}_D(\mathbf{x}, t) &= \frac{1}{c}\int \mathrm{d}^3\mathbf{x}'' \frac{\mathbf{j}_t(\mathbf{x}'', t)}{\|\mathbf{x} - \mathbf{x}''\|} \\ &= \frac{1}{4\pi c}\int \mathrm{d}^3\mathbf{x}'' \frac{1}{\|\mathbf{x} - \mathbf{x}''\|}\left[\nabla'' \times \left(\nabla'' \times \int d^3x' \frac{\mathbf{j}(\mathbf{x}', t)}{\|\mathbf{x}'' - \mathbf{x}'\|}\right)\right].\end{aligned} \tag{4.17}$$

This expression, however, is rather unwieldy, and in most calculations it is preferable to use the more direct expression $\mathbf{A}_C^{qs}$ to express the vector potential of the Darwin system.

At this point we now have two different 'Darwin' vector potentials. By calculating the instantaneous limit of the Coulomb gauge equations, we arrived at the form $\mathbf{A}_C^{qs}$ given in (4.9). On the other hand, by taking the non-retarded limit of the field equations themselves, we obtained the above expression for $\mathbf{A}_D$. The glaring question, then, is are the two equal? This is the subject of the next section, in which we attempt to show by direct calculation the two forms of the potential are indeed the same.

## 4.3 Equivalence of $\mathbf{A}_C^{qs}$ and $\mathbf{A}_D$

We now have two expressions for the vector potential of the Darwin system: the quasistatic form $\mathbf{A}_C^{qs}$, obtained as an approximation to the retarded solution of the Coulomb gauge equations and given by (4.9); and the 'Darwin' form $\mathbf{A}_D$, obtained as the exact solution to approximated field equations and given by (4.17). Looking at these expressions, it is not at all clear that the two potentials are the same. Of course we hope they are, for if not, then we have two different, inequivalent formulations of the Darwin system. We must therefore find a way of transforming $\mathbf{A}_D$ into $\mathbf{A}_C^{qs}$ to establish clearly that the two formulations do in fact describe the same physics.

To this end, we start with $\mathbf{A}_D$ as given in (4.17) and write the cross products in index notation:

$$\mathbf{A}_D = \frac{1}{4\pi c}\int \mathrm{d}^3\mathbf{x}''\,\mathrm{d}^3\mathbf{x}'\partial_j''\partial_l''\left(\frac{1}{\|\mathbf{x}-\mathbf{x}''\|}\right)\frac{\epsilon_{ijk}\epsilon_{klm}\mathrm{j}_m'}{\|\mathbf{x}''-\mathbf{x}'\|}, \tag{4.18}$$

where $\partial_j'' = \partial/\partial x_j''$ and $\mathbf{j}' = \mathbf{j}(\mathbf{x}', t)$. Using the identity $\epsilon_{kij}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$, this becomes

$$\mathbf{A}_D = \frac{1}{4\pi c}\int \frac{\mathrm{d}^3\mathbf{x}''\,\mathrm{d}^3\mathbf{x}'}{\|\mathbf{x}''-\mathbf{x}'\|}\left[\mathrm{j}'_m\left(\partial''_m\partial''_i\frac{1}{\|\mathbf{x}-\mathbf{x}''\|}\right)-\mathrm{j}'_i\left(\partial''_l\partial''_l\frac{1}{\|\mathbf{x}-\mathbf{x}''\|}\right)\right]$$

$$= \frac{1}{4\pi c}\int \frac{\mathrm{d}^3\mathbf{x}''\,\mathrm{d}^3\mathbf{x}'}{\|\mathbf{x}''-\mathbf{x}'\|}(\mathbf{j}'\cdot\nabla'')\nabla''\frac{1}{\|\mathbf{x}-\mathbf{x}''\|}+\frac{1}{c}\int\frac{\mathrm{d}^3\mathbf{x}''\,\mathrm{d}^3\mathbf{x}'\mathbf{j}'}{\|\mathbf{x}''-\mathbf{x}'\|}\delta^{(3)}\left(\mathbf{x}-\mathbf{x}''\right). \tag{4.19}$$

Integration over the delta function gives

$$\mathbf{A}_D = \frac{1}{c}\int\frac{\mathrm{d}^3\mathbf{x}'\mathbf{j}(\mathbf{x}',t)}{\|\mathbf{x}-\mathbf{x}'\|}+\frac{1}{4\pi c}\int\frac{\mathrm{d}^3\mathbf{x}''\,\mathrm{d}^3\mathbf{x}'}{\|\mathbf{x}''-\mathbf{x}'\|}(\mathbf{j}'\cdot\nabla'')\nabla''\frac{1}{\|\mathbf{x}-\mathbf{x}''\|}. \tag{4.20}$$

If we now use the fact that

$$\frac{\partial}{\partial x''_i}\frac{1}{\|\mathbf{x}-\mathbf{x}''\|} = -\frac{\partial}{\partial x_i}\frac{1}{\|\mathbf{x}-\mathbf{x}''\|}, \tag{4.21}$$

then we may rewrite $\mathbf{A}_D$ as

$$\mathbf{A}_D = \frac{1}{c}\int\frac{\mathrm{d}^3\mathbf{x}'\mathbf{j}(\mathbf{x}',t)}{\|\mathbf{x}-\mathbf{x}'\|}+\frac{1}{4\pi c}\int\mathrm{d}^3\mathbf{x}'(\mathbf{j}'\cdot\nabla)\nabla\int\frac{\mathrm{d}^3\mathbf{x}''}{\|\mathbf{x}''-\mathbf{x}'\|\,\|\mathbf{x}-\mathbf{x}''\|}. \tag{4.22}$$

The last term now contains the Darwin integral (*cf.* (D.13)):

$$\int\frac{\mathrm{d}^3\mathbf{x}}{\|\mathbf{x}-\mathbf{a}\|\,\|\mathbf{x}-\mathbf{b}\|} = -2\pi\,\|\mathbf{a}-\mathbf{b}\|\ . \tag{4.23}$$

Using this result, we have

$$\mathbf{A}_D = \frac{1}{c}\int\frac{\mathrm{d}^3\mathbf{x}'\mathbf{j}(\mathbf{x}',t)}{\|\mathbf{x}-\mathbf{x}'\|}-\frac{1}{2c}\int\mathrm{d}^3\mathbf{x}'(\mathbf{j}'\cdot\nabla)\nabla\,\|\mathbf{x}-\mathbf{x}'\|$$

$$= \frac{1}{c}\int\frac{\mathrm{d}^3\mathbf{x}'\mathbf{j}(\mathbf{x}',t)}{\|\mathbf{x}-\mathbf{x}'\|}-\frac{1}{2c}\int\mathrm{d}^3\mathbf{x}'\frac{\mathbf{j}'-(\mathbf{j}'\cdot\hat{\mathbf{r}}')\hat{\mathbf{r}}'}{\|\mathbf{x}-\mathbf{x}'\|}, \tag{4.24}$$

so that finally

$$\mathbf{A}_D(\mathbf{x},t) = \frac{1}{2c}\int\mathrm{d}^3\mathbf{x}'\frac{\mathbf{j}(\mathbf{x}',t)+\hat{\mathbf{r}}'(\hat{\mathbf{r}}'\cdot\mathbf{j}(\mathbf{x}',t))}{\|\mathbf{x}-\mathbf{x}'\|} = \mathbf{A}_C^{qs}(\mathbf{x},t). \tag{4.25}$$

The two forms of the vector potential are equal, and therefore the two different methods of arriving at the Darwin system are physically equivalent. This argument rests crucially on the value of the Darwin integral (D.13); although computation of this integral here would take the discussion too far afield, the seemingly contradictory nature of its solution involves subtle argumentation, and this is given in the appendix.

## 4.4 Variational Generalia

Suppose we are given an action for fields $\phi$, $\mathbf{A}$ and sources $\mathbf{q}$ of the form

$$S[\mathbf{q}, \phi, \mathbf{A}] = S_{\text{kinetic}}[\mathbf{q}] + S_{\text{field}}[\phi, \mathbf{A}] + S_{\text{coupling}}[\mathbf{q}, \phi, \mathbf{A}], \tag{4.26}$$

where

$$S_{\text{coupling}}[\mathbf{q}, \phi, \mathbf{A}] = \int \mathrm{d}t \, \mathrm{d}^3\mathbf{x} \left[ -\rho\phi + \frac{\mathbf{j}}{c} \cdot \mathbf{A} \right], \tag{4.27}$$

and where $\rho$, $\mathbf{j}$ are given in terms of $\mathbf{q}$. Variation with respect to $\phi$ and $\mathbf{A}$ link these fields to the sources $\rho, \mathbf{j}$:

$$O_1[\phi] = \rho, \qquad O_2[\mathbf{A}] = \mathbf{j}, \tag{4.28}$$

where in general $O_{1,2}$ are linear integro-differential operators which depend on $\phi$ and $\mathbf{A}$ both. Variations with respect to $\mathbf{q}$ in turn link the acceleration to the fields:

$$\text{accel.} = -F(\phi, \mathbf{A}) := -\left[ -\frac{\partial \rho}{\partial \mathbf{q}} \phi + \frac{1}{c} \frac{\partial \mathbf{j}}{\partial \mathbf{q}} \cdot \mathbf{A} \right], \tag{4.29}$$

where "accel." is the variation of $S_{\text{kinetic}}$. If we can solve the equations for $\phi$ and $\mathbf{A}$ in terms of $\rho$ and $\mathbf{j}$, respectively, then we may write

$$\phi_s(\mathbf{x}) = \int K_\phi\left(\mathbf{x} \mid \mathbf{x}'\right)\rho(\mathbf{x}'), \qquad \mathbf{A}_s(\mathbf{x}) = \int K_{\mathbf{A}}\left(\mathbf{x} \mid \mathbf{x}'\right)\mathbf{j}(\mathbf{x}'). \tag{4.30}$$

Now, we may substitute $\phi_s$ and $\mathbf{A}_s$ in $S_{\text{field}}$ and $S_{\text{coupling}}$ in order to write the action $S$ solely in terms of $\mathbf{q}$:

$$S_{\text{new}}[\mathbf{q}] = S[\mathbf{q}, \phi_s(\mathbf{q}), \mathbf{A}_s(\mathbf{q})]. \tag{4.31}$$

This procedure is completely general in the sense that we have not specified the nature of $\mathbf{q}$. Inserting these into $\rho$ and $\mathbf{j}$, then writing $\phi$ and $\mathbf{A}$ in terms of $\rho$ and $\mathbf{j}$, gives $S$ in terms of $\mathbf{q}$ alone.

We can illustrate the process if we now take as our field action the electromagnetic action

$$\begin{aligned} S_{EM} &= \frac{1}{8\pi}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}(E^2 - B^2) \\ &= \frac{1}{8\pi}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\left[\mathbf{A}\cdot\left(\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2}\right)\right. \\ &\qquad \left. + \frac{2}{c}\frac{\partial\mathbf{A}}{\partial t}\cdot\nabla\phi + \left\|\nabla\phi\right\|^2 - \mathbf{A}\cdot\nabla(\nabla\cdot\mathbf{A})\right], \end{aligned} \tag{4.32}$$

we may use the general equations for the potentials, namely

$$\begin{aligned} \nabla^2\phi + \frac{1}{c}\frac{\partial}{\partial t}\nabla\cdot\mathbf{A} &= -4\pi\rho, \\ \nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} - \nabla\left(\nabla\cdot\mathbf{A} + \frac{1}{c}\frac{\partial\phi}{\partial t}\right) &= -\frac{4\pi}{c}\mathbf{j}, \end{aligned} \tag{4.33}$$

to recast the form of the field action. Inserting the $\mathbf{j}$ equation, we arrive at the expression

$$S_{EM} = \frac{1}{8\pi} \int \mathrm{d}t \, \mathrm{d}^3\mathbf{x} \left[ \mathbf{A} \cdot \left( -\frac{4\pi}{c} \mathbf{j} \right) - \mathbf{A} \cdot \frac{1}{c} \frac{\partial \nabla \phi}{\partial t} - \phi \nabla^2 \phi \right]. \tag{4.34}$$

We may then integrate by parts once more and insert the $\rho$ equation to find

$$S_{EM} = -\frac{1}{2} \int \mathrm{d}t \, \mathrm{d}^3\mathbf{x} \left[ -\rho\phi + \frac{\mathbf{j}}{c} \cdot \mathbf{A} \right], \tag{4.35}$$

so that the field term combines with the source coupling term to give

$$S_{EM} + \int \mathrm{d}t \, \mathrm{d}^3\mathbf{x} \left[ -\rho\phi + \frac{\mathbf{j}}{c} \cdot \mathbf{A} \right] = \frac{1}{2} \int \mathrm{d}t \, \mathrm{d}^3\mathbf{x} \left[ -\rho\phi + \frac{\mathbf{j}}{c} \cdot \mathbf{A} \right]. \tag{4.36}$$

This result is completely general and relies only on the form of the electromagnetic field action and the general equations which the potentials satisfy.

Returning to the abstract arguments, let us merely assume the operators in (4.28) relating the potentials to the sources are *linear*:

$$\phi = L_1[\rho], \qquad \mathbf{A} = L_2[\mathbf{j}]. \tag{4.37}$$

This allows us to write

$$S_{\text{new}}[\mathbf{q}] = S[\mathbf{q}, \phi_s, \mathbf{A}_s] = S[\mathbf{q}, L_1[\rho(\mathbf{q})], L_2[\mathbf{j}(\mathbf{q})]], \tag{4.38}$$

and we may thus compute the variation of the above combination of the coupling and field terms. Focusing on the $\rho\phi$ term, we find the variations with respect to $\mathbf{q}$:

$$\begin{aligned}
\delta\,\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}(-\rho\phi_s) &= \delta\,\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}(-\rho L_1[\rho]) \\
&= \frac{d}{d\epsilon}\bigg|_{\epsilon=0}\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}[-(\rho+\epsilon\delta\rho)L_1[\rho+\epsilon\delta\rho]] \\
&= \frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}[-\delta\rho L_1[\rho]-\rho L_1[\delta\rho]] \\
&= -\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\delta\rho(L_1+L_1^\dagger)[\rho],
\end{aligned} \tag{4.39}$$

where $L_1^\dagger$ is the adjoint of $L_1$. If we further assume that $L_1$ is self-adjoint, so that $L_1^\dagger = L_1$, we find that

$$\delta\,\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}(-\rho\phi_s) = -\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\delta\rho\cdot 2L_1[\rho] = \int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}[-\delta\rho\;\phi_s], \tag{4.40}$$

and the factor of $1/2$ is cancelled. An analogous calculation holds for the $\mathbf{j}\cdot\mathbf{A}$ term, and so $\mathbf{q}$-variation of the action $S_{\text{new}}[\mathbf{q}]$ yields

$$\text{accel.} = -F(\phi_s,\mathbf{A}_s) = -\left[-\frac{\partial\rho}{\partial\mathbf{q}}\phi_s + \frac{1}{c}\frac{\partial\mathbf{j}}{\partial\mathbf{q}}\cdot\mathbf{A}_s\right] \tag{4.41}$$

as before, except that here $\phi_s$ and $\mathbf{A}_s$ are shorthand for (4.37). In particular, the right hand side is a function of $\mathbf{q}$ and its derivatives alone.

## 4.5 The Darwin Field Action

Starting from the electromagnetic action, we may verify that dropping terms of order greater than $(1/c)^2$ leads to the proper Darwin field equations. Starting with the electromagnetic field and coupling terms, we may expand the squared terms to find

$$
\begin{aligned}
S_M[\phi, \mathbf{A}] = & \int \mathrm{d}t\, \mathrm{d}^3\mathbf{x} \left[ -\rho\phi + \frac{\mathbf{j}}{c} \cdot \mathbf{A} \right] \\
& + \frac{1}{8\pi} \int \mathrm{d}t\, \mathrm{d}^3\mathbf{x} \left[ \left\| \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} + \nabla\phi \right\|^2 - \|\nabla \times \mathbf{A}\|^2 \right] \\
= & \int \mathrm{d}t\, \mathrm{d}^3\mathbf{x} \left[ -\rho\phi + \frac{\mathbf{j}}{c} \cdot \mathbf{A} + \frac{1}{8\pi} \left( \frac{1}{c^2} \left( \frac{\partial \mathbf{A}}{\partial t} \right)^2 + \frac{2}{c} \frac{\partial \mathbf{A}}{\partial t} \cdot \nabla\phi + \|\nabla\phi\|^2 \right.\right. \\
& \left.\left. - \mathbf{A} \cdot \left( \nabla \left( \nabla \cdot \mathbf{A} \right) - \nabla^2 \mathbf{A} \right) \right) \right], \qquad (4.42)
\end{aligned}
$$

where an integration by parts has been performed to write $(\nabla \times \mathbf{A})^2$ as $\mathbf{A} \cdot (\nabla \times \nabla \times \mathbf{A})$, and then the identity $\nabla \times \nabla \times = \nabla(\nabla \cdot -\nabla^2$ has been applied. ("$M$" here stands for "Maxwell".)

Note however that the term $(1/c^2)(\partial \mathbf{A}/\partial t)^2$ is higher order than what is consistent with the Darwin approximation. Specifically, the vector potential itself scales as $\mathbf{A} \sim 1/c$. This term then is of order $1/c^4$, and is therefore well beyond the $1/c^2$ order of the Darwin approximation. We may therefore omit this term from the action, and work instead with the following expression for the Darwin fields:

$$
\begin{aligned}
S_{\mathrm{df}}[\phi, \mathbf{A}] = & \int \mathrm{d}t\, \mathrm{d}^3\mathbf{x} \left[ -\rho\phi + \frac{\mathbf{j}}{c} \cdot \mathbf{A} \right] \\
& + \frac{1}{8\pi} \int \mathrm{d}t\, \mathrm{d}^3\mathbf{x} \left[ \frac{2}{c} \frac{\partial \mathbf{A}}{\partial t} \cdot \nabla\phi + \|\nabla\phi\|^2 - \mathbf{A} \cdot \left( \nabla \left( \nabla \cdot \mathbf{A} \right) - \nabla^2 \mathbf{A} \right) \right].
\end{aligned}
\qquad (4.43)
$$

Variation with respect to $\phi$ yields

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S_{\mathrm{df}}[\phi+\epsilon\delta\phi,\mathbf{A}] = \int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\;\delta\phi\;\left[-\rho - \frac{1}{4\pi}\left(\frac{1}{c}\frac{\partial}{\partial t}\nabla\cdot\mathbf{A} + \nabla^2\phi\right)\right], \tag{4.44}$$

so that one equation of motion is

$$\nabla^2\phi + \frac{1}{c}\frac{\partial}{\partial t}\nabla\cdot\mathbf{A} = -4\pi\rho. \tag{4.45}$$

Variation with respect to $\mathbf{A}$ yields

$$\begin{aligned}\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S_{\mathrm{df}}[\phi,\mathbf{A}+\epsilon\delta\mathbf{A}] = \int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\,\Bigg[&\frac{\mathbf{j}}{c}\cdot\delta\mathbf{A} + \frac{1}{8\pi}\left(\frac{2}{c}\nabla\phi\cdot\frac{\partial\delta\mathbf{A}}{\partial t}\right)\\ &-\frac{1}{8\pi}\delta\mathbf{A}\cdot\big(\nabla\left(\nabla\cdot\mathbf{A}\right) - \nabla^2\mathbf{A}\big)\\ &-\frac{1}{8\pi}\mathbf{A}\cdot\big(\nabla\left(\nabla\cdot\delta\mathbf{A}\right) - \nabla^2\delta\mathbf{A}\big)\Bigg].\end{aligned} \tag{4.46}$$

Integration by parts gives

$$\delta S_{\mathrm{df}} = \int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\delta\mathbf{A}\cdot\left[\frac{\mathbf{j}}{c} - \frac{1}{4\pi}\left(\frac{1}{c}\frac{\partial\nabla\phi}{\partial t} + \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A}\right)\right], \tag{4.47}$$

so that the other equation of motion is

$$-\nabla^2\mathbf{A} + \nabla\left(\nabla\cdot\mathbf{A} + \frac{1}{c}\frac{\partial\phi}{\partial t}\right) = \frac{4\pi}{c}\mathbf{j}. \tag{4.48}$$

The resulting equations are equivalent to the Darwin field equations

$$\begin{aligned}\nabla^2\phi &= -4\pi\rho,\\ \nabla^2\mathbf{A} &= -\frac{4\pi}{c}\left(\mathbf{j} - \frac{1}{4\pi}\frac{\partial\nabla\phi}{\partial t}\right)\end{aligned} \tag{4.49}$$

provided we choose Coulomb gauge $\nabla \cdot \mathbf{A} = 0$.

We therefore see that starting with the standard electromagnetic action and dropping the term $(1/c^2)(\partial \mathbf{A}/\partial t)^2$ is consistent with the Darwin field equations when we choose Coulomb gauge. We might naively summarize this by writing a Darwin field action in the form

$$S_{\text{df}}[\phi, \mathbf{A}, \lambda] = \int \mathrm{d}t\, \mathrm{d}^3\mathbf{x} \left[-\rho\phi + \frac{\mathbf{j}}{c} \cdot \mathbf{A}\right] + \frac{1}{8\pi} \int \mathrm{d}t\, \mathrm{d}^3\mathbf{x} \left[E^2 - B^2\right]$$
$$+ \frac{1}{8\pi} \int \mathrm{d}t\, \mathrm{d}^3\mathbf{x} \left[\lambda \nabla \cdot \mathbf{A} - \frac{1}{c^2} \left(\frac{\partial \mathbf{A}}{\partial t}\right)^2\right], \tag{4.50}$$

and varying with respect to $\phi$, $\mathbf{A}$, and the Lagrange multiplier $\lambda$. It is clear from this that the Darwin system is *not* a gauge-invariant theory. In fact, the stipulation that one work in Coulomb gauge is closely tied to local charge conservation.

## 4.6 Darwin and Local Charge Conservation

Let us return to the more general equations (4.4) derived from the electromagnetic field action, but with the term $(1/c^2)(\partial \mathbf{A}/\partial t)^2$ removed:

$$\nabla^2 \phi + \frac{1}{c} \frac{\partial}{\partial t} \nabla \cdot \mathbf{A} = -4\pi\rho,$$
$$-\nabla^2 \mathbf{A} + \nabla \left(\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \phi}{\partial t}\right) = \frac{4\pi}{c} \mathbf{j}. \tag{4.51}$$

If we take the time-derivative of the first and divergence of the second, we find that

$$4\pi \left(\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j}\right) = -\frac{1}{c} \frac{\partial^2}{\partial t^2} \nabla \cdot \mathbf{A}. \tag{4.52}$$

Since $\mathbf{A} \sim 1/c$, then $(1/c)\partial^2(\nabla \cdot \mathbf{A})/\partial t^2 \sim 1/c^2$, and so this term must be retained in general to the order at which we are working. But to have local charge conservation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \qquad \Longrightarrow \qquad \frac{\partial^2}{\partial t^2} \nabla \cdot \mathbf{A} = 0. \tag{4.53}$$

In general this means that $\nabla \cdot \mathbf{A}$ must be a linear function of time: $\nabla \cdot \mathbf{A} = f(\mathbf{x})t + g(\mathbf{x})$. We may suppose that linear $t$-dependence is unphysical, which leaves $\nabla \cdot \mathbf{A} = g(\mathbf{x})$. There is no obvious physical argument why we must have $g(\mathbf{x}) \equiv 0$, but this makes clear that the choice of Coulomb gauge $\nabla \cdot \mathbf{A} = 0$ is the only natural gauge choice for the Darwin system that maintains local charge conservation.

We may note in passing that the Darwin equation for the vector potential is often written

$$\nabla^2 \mathbf{A} = -\frac{4\pi}{c} \mathbf{j}_t, \tag{4.54}$$

where $\mathbf{j}_t$ is the transverse component of the current $\mathbf{j}$. This requires writing

$$\mathbf{j}_l = \frac{1}{4\pi} \frac{\partial \nabla \phi}{\partial t}, \tag{4.55}$$

for the longitudinal component of $\mathbf{j}$. Such an identification, however, *explicitly* uses local charge conservation to write replace $\nabla \cdot \mathbf{j}$ with $-\partial \rho / \partial t$. The Darwin field equations therefore rely on local charge conservation, requiring the imposition of Coulomb gauge $\nabla \cdot \mathbf{A} = 0$ as a constraint additional to dropping the term $(1/c^2)(\partial \mathbf{A}/\partial t)^2$ from the action.

# Chapter 5

# Self-Consistent Darwin Theory

In Section 3.1, particularly equation (3.19), we arrived at an expression for the Darwin interaction of charged particles starting from an expression in which the potentials $\phi$ and $\mathbf{A}$ were given. That these fields should satisfy Maxwell's equations up to second order in $v/c$ must be added as a constraint within this framework. There already exists, however, a well-known action formulation whereby variation on $\phi$ and $\mathbf{A}$ yields Maxwell's equations directly, namely

$$S_{EM}[\phi, \mathbf{A}] := \frac{1}{8\pi} \int \mathrm{d}t \, \mathrm{d}^3\mathbf{x} \left[ \left( -\frac{1}{c} \frac{\partial \mathbf{A}(\mathbf{x}, t)}{\partial t} - \nabla \phi(\mathbf{x}, t) \right)^2 - (\nabla \times \mathbf{A}(\mathbf{x}, t))^2 \right]. \tag{5.1}$$

It is natural, therefore, to investigate whether a sensible theory arises when one takes as the starting point a total action which is a functional of both particles and fields, together with an interaction term, and then follows the Darwin procedure used previously.

## 5.1 Particle Theory

In our endeavor to find a Darwin particle theory whose fields are self-consitently governed by "approximated" Maxwell's equations, we start with the well-known action coupling charged particles to electromagnetic fields:

$$
\begin{aligned}
&S[\mathbf{q}_a;\phi,\mathbf{A}]\\
&:= \; -\sum_a m_a c^2 \int \mathrm{d}t \sqrt{1-\frac{\dot{q}_a^2(t)}{c^2}}\\
&\quad +\sum_a e_a \int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\,\delta^{(3)}\left(\mathbf{x}-\mathbf{q}_a(t)\right)\left[-\phi(\mathbf{x},t)+\frac{1}{c}\mathbf{A}(\mathbf{x},t)\cdot\dot{\mathbf{q}}_a\right]\\
&\quad +\frac{1}{8\pi}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\left[\left(-\frac{1}{c}\frac{\partial\mathbf{A}(\mathbf{x},t)}{\partial t}-\nabla\phi(\mathbf{x},t)\right)^2-\left(\nabla\times\mathbf{A}(\mathbf{x},t)\right)^2\right].
\end{aligned}
\tag{5.2}
$$

We want to expand to the only consistent Darwin action — we expand everything, dropping the term $(1/c^2)(\partial\mathbf{A}/\partial t)^2$ and other terms of order higher than $(v/c)^2$. We expand the kinetic term as

$$
S_K := -\sum_a m_a c^2 \int \mathrm{d}t\sqrt{1-\frac{\dot{q}_a^2}{c^2}} \sim \sum_a \int \mathrm{d}t\left(\frac{m_a\dot{q}_a^2}{2}+\frac{m_a\dot{q}_a^4}{8c^2}\right). \tag{5.3}
$$

We expand the field term and drop the term $(1/c^2)(\partial\mathbf{A}/\partial t)^2$ for reasons mentioned in Section 4.5:

$$
\begin{aligned}
S_F &:= \frac{1}{8\pi}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\left[\left(\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}+\nabla\phi\right)^2-(\nabla\times\mathbf{A})^2\right]\\
&\approx \frac{1}{8\pi}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\left[-\phi\frac{1}{c}\frac{\partial}{\partial t}\nabla\cdot\mathbf{A}-\frac{1}{c}\mathbf{A}\cdot\frac{\partial\nabla\phi}{\partial t}-\phi\nabla^2\phi\right]\\
&\qquad -\frac{1}{8\pi}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\left[\mathbf{A}\cdot\left(\nabla(\nabla\cdot\mathbf{A})-\nabla^2\mathbf{A}\right)\right],
\end{aligned}
\tag{5.4}
$$

where the term $(2/c)(\partial\mathbf{A}/\partial t)\cdot\nabla\phi$ has been written as the sum of two similar terms, one of which is then integrated by parts in $\mathbf{x}$, the other in $t$. We may now insert the Darwin field equations,

$$\nabla^2\phi_D = -4\pi\rho,$$
$$\nabla^2\mathbf{A}_D = -\frac{4\pi}{c}\left(\mathbf{j} - \frac{1}{4\pi}\frac{\partial\nabla\phi_D}{\partial t}\right), \tag{5.5}$$

expressing the fields in terms of sources. Using $\nabla\cdot\mathbf{A} = 0$, we can express $S_F$ as a functional of $\mathbf{q}$

$$S_F = \frac{1}{2}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\left[\rho\phi_D - \frac{\mathbf{j}}{c}\cdot\mathbf{A}_D\right], \tag{5.6}$$

where

$$\rho(\mathbf{x},t) = \sum_a e_a\delta^{(3)}\left(\mathbf{x} - \mathbf{q}_a(t)\right),$$
$$\mathbf{j}(\mathbf{x},t) = \sum_a e_a\dot{\mathbf{q}}_a(t)\delta^{(3)}\left(\mathbf{x} - \mathbf{q}_a(t)\right). \tag{5.7}$$

(This can be read in either direction — *i.e.* we could have started with a completely general action whose coupling term was $-\rho\phi + \mathbf{A}\cdot\mathbf{j}/c$, rather than delta functions.) Then we can combine the field and coupling terms in the action to arrive at the following:

$$\begin{aligned} S_D[\mathbf{q}] = \sum_a \int \mathrm{d}t\left(\frac{m_a\dot{q}_a^2}{2} + \frac{m_a\dot{q}_a^4}{8c^2}\right) \\ + \frac{1}{2}\sum_a e_a\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\delta^{(3)}\left(\mathbf{x} - \mathbf{q}_a\right)\left[-\phi_D + \frac{\dot{\mathbf{q}}_a}{c}\cdot\mathbf{A}_D\right], \end{aligned} \tag{5.8}$$

where $\phi_D$ and $\mathbf{A}_D$ are solutions of the Darwin field equations. We may write this more generally as

$$S_D[\mathbf{q}] = \sum_a \int \mathrm{d}t \left( \frac{m_a \dot{q}_a^2}{2} + \frac{m_a \dot{q}_a^4}{8c^2} \right)$$
$$+ \frac{1}{2} \int \mathrm{d}t\, \mathrm{d}^3\mathbf{x} \left[ -\rho\phi_D + \frac{\mathbf{j}}{c} \cdot \mathbf{A}_D \right], \tag{5.9}$$

showing that none of the calculations so far have been specific to particles, but hold just as well for a continuum.

We must now find expressions for $\phi_D$ and $\mathbf{A}_D$. These of course come from solving the Darwin field equations themselves. Using the expression for the density in terms of delta functions centered on the particles, the scalar potential is simple to compute by using the Green's function for Poisson's equation:

$$\phi_D(\mathbf{x}, t) = \sum_b e_b \int \mathrm{d}^3\mathbf{x}' \frac{\delta^{(3)}\left(\mathbf{x}' - \mathbf{q}_b\right)}{\|\mathbf{x}' - \mathbf{x}\|} = \sum_b \frac{e_b}{\|\mathbf{x} - \mathbf{q}_b\|}. \tag{5.10}$$

The method for solving the equation for the vector potential is actually the same, once the equation is expressed in terms of the transverse current. To this end, we write

$$\nabla^2 \mathbf{A}_D = -\frac{4\pi}{c}\mathbf{j}_t = -\frac{1}{c}\nabla \times \left( \nabla \times \int \mathrm{d}^3\mathbf{x}' \frac{\mathbf{j}(\mathbf{x}', t)}{\|\mathbf{x} - \mathbf{x}'\|} \right). \tag{5.11}$$

Again solution amounts to inverting Poisson's equation, and so we have

$$\mathbf{A}_D(\mathbf{x}, t) = \frac{1}{4\pi c} \int \mathrm{d}^3\mathbf{x}'' \frac{1}{\|\mathbf{x} - \mathbf{x}''\|} \left[ \nabla'' \times \left( \nabla'' \times \int \mathrm{d}^3\mathbf{x}' \frac{\mathbf{j}(\mathbf{x}', t)}{\|\mathbf{x}'' - \mathbf{x}'\|} \right) \right]. \tag{5.12}$$

The form, however, is rather inconvenient for our purposes, and so we recast it as

$$\mathbf{A}_D(\mathbf{x},t) = \frac{1}{2c}\int \mathrm{d}^3\mathbf{x}'\frac{\mathbf{j}(\mathbf{x}',t) + (\mathbf{j}(\mathbf{x}',t)\cdot\hat{\mathbf{r}}')\hat{\mathbf{r}}'}{\|\mathbf{x}-\mathbf{x}'\|} = \mathbf{A}_C^{qs}(\mathbf{x},t). \tag{5.13}$$

Writing the current now as a sum of delta functions centered on the particles, this gives

$$\begin{aligned}\mathbf{A}_D(\mathbf{x},t) &= \frac{1}{2c}\sum_b e_b \int \mathrm{d}^3\mathbf{x}'\delta^{(3)}\left(\mathbf{x}'-\mathbf{q}_b\right)\frac{\dot{\mathbf{q}}_b + (\dot{\mathbf{q}}_b\cdot\hat{\mathbf{r}}')\hat{\mathbf{r}}'}{\|\mathbf{x}-\mathbf{x}'\|} \\ &= \frac{1}{2c}\sum_b e_b\frac{\dot{\mathbf{q}}_b + (\dot{\mathbf{q}}_b\cdot\hat{\mathbf{r}}_b)\hat{\mathbf{r}}_b}{\|\mathbf{x}-\mathbf{q}_b\|}. \end{aligned} \tag{5.14}$$

Inserting the expressions for $\phi_D$ and $\mathbf{A}_D$ into the action, we find the following:

$$\begin{aligned}S_D[\mathbf{q}_a] = \sum_a \int \mathrm{d}t\left(\frac{m_a\dot{q}_a^2}{2} + \frac{m_a\dot{q}_a^4}{8c^2}\right) + \frac{1}{2}\sum_{a,b}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\frac{-e_ae_b\delta^{(3)}\left(\mathbf{x}-\mathbf{q}_a\right)}{\|\mathbf{x}-\mathbf{q}_b\|} \\ +\frac{1}{2}\sum_{a,b}\int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\frac{e_ae_b\delta^{(3)}\left(\mathbf{x}-\mathbf{q}_a\right)}{\|\mathbf{x}-\mathbf{q}_b\|}\frac{1}{2c^2}\left[\dot{\mathbf{q}}_a\cdot\dot{\mathbf{q}}_b + \left(\dot{\mathbf{q}}_a\cdot\hat{\mathbf{r}}_b\right)\left(\dot{\mathbf{q}}_b\cdot\hat{\mathbf{r}}_b\right)\right].\end{aligned} \tag{5.15}$$

We may then perform the integrations over the delta functions to obtain an action solely in terms of particle positions an velocities. We of course remove by hand infinite contributions which arise when $a = b$. The final result is

$$S_D[\mathbf{q}_a] = \sum_a \int \mathrm{d}t \left( \frac{m_a \dot{q}_a^2}{2} + \frac{m_a \dot{q}_a^4}{8c^2} \right) + \frac{1}{2} \sum_{\substack{a,b \\ b \neq a}} \int \mathrm{d}t \frac{-e_a e_b}{\|\mathbf{q}_a - \mathbf{q}_b\|}$$

$$+ \frac{1}{2} \sum_{\substack{a,b \\ b \neq a}} \int \mathrm{d}t \frac{e_a e_b}{2c^2 \|\mathbf{q}_a - \mathbf{q}_b\|} \left[ \dot{\mathbf{q}}_a \cdot \dot{\mathbf{q}}_b + (\dot{\mathbf{q}}_a \cdot \hat{\mathbf{r}}_{ab}) (\dot{\mathbf{q}}_b \cdot \hat{\mathbf{r}}_{ab}) \right]. \tag{5.16}$$

This agrees with expressions in the literature.

## 5.2 Epilogue

At this point it is a matter of some small interest to compare the methods of this chapter with investigations carried out by Wheeler and Feynman. In the 1940s, Wheeler and Feynman [56, 57] published a series of papers studying the formulation of classical electromagnetism as an action-at-a-distance theory. In the first paper they established that taking one-half the sum or difference of the retarded and advanced electromagnetic fields of absorbers or radiator led to the proper field expressions of the radiative reaction force. They subsequently generalized the procedure in the attempt to reproduce proper field equations in general solely from consideration of the source charges alone.[1]

The method of this chapter has been to start with fields $\phi$, $\mathbf{A}$ and particles $\mathbf{q}$, each obeying the proper respective equations of motion, and then

[1]In addition to the original papers, see discussions in Rohrlich [50] and Panofsky and Phillips [48].

to formulate the theory in terms of particles alone. This occurs by substituting for the fields in terms of the source charges (equations (5.10) and (5.14)), and then by inserting the resulting relations back into the Lagrangian. This yields a particle action, with no reference to fields.

If we now read the chapter in reverse, we see that this is another small illustration of Wheeler and Feynman's procedure. If we start from the purely particle action $S_D[\mathbf{q}_a]$ given in (5.16), and then *define* the fields $\phi$ and $\mathbf{A}$ according to (5.10) and (5.14) respectively, the arguments of this chapter taken in reverse order show that the resulting fields obey the Darwin field equations. Thus we have achieved an action-at-a-distance Darwin theory akin to the theory of Wheeler-Feynman. This is naturally to be expected, since Darwin himself developed the theory for the express purpose of having an action-at-a-distance formulation, but the observation is general. Whenever we pursue particle theories along the lines of this chapter, we may look to see if the process in reverse leads to a Wheeler-Feynman theory of the original system.

# Chapter 6

# A Low-Darwin System

## 6.1 Introduction

In 1958 F. E. Low published a short article in which he derived a variational principle for the Vlasov equation for a charged gas in an electromagnetic field [40]. As a starting point, he assumed an initial distribution of particles $f_0(z_0)$ was given at time $t = 0$, where $z_0 = (\mathbf{x}_0, \mathbf{v}_0)$. The particles move on trajectories $\mathbf{q}(z, t)$ such that $\mathbf{q}(z_0, 0) = \mathbf{x}_0$ and $\dot{\mathbf{q}}(z_0, 0) = \mathbf{v}_0$. The distribution function is constant on particle trajectories, so that

$$f(\mathbf{q}, \dot{\mathbf{q}}, t) = f_0(z_0). \tag{6.1}$$

We may recast this condition as

$$\frac{df}{dt} = \frac{\partial f(z,t)}{\partial t} + \dot{\mathbf{q}} \cdot \frac{\partial f(z,t)}{\partial \mathbf{x}} + \ddot{\mathbf{q}} \cdot \frac{\partial f(z,t)}{\partial \mathbf{v}} = 0 \tag{6.2}$$

along the trajectories, which is the Vlasov equation. Given the usefulness of the Darwin system in plasma simulation, it is natural to seek a formulation for the Darwin system in terms of the prescription of Low. This is the focus of the present chapter.

The actual method of this chapter follows the presentation of Ye and Morrison [58]. This method amounts to expressing the potentials in terms of kernels of the associated equations, and then exploiting the symmetries of the kernel in the variations. For more on the methods of variation, consult Morrison's review article on Hamiltonian methods applied to the ideal fluid [43].

Many of the calculations of this chapter have already been treated in a more general setting. But because of the novelty of the Low formalism, the calculations are presented again in order to gain a better understanding of how the distribution function $f$ enters into the scheme. The form of the Vlasov equation is taken as a given; the task is to describe the equations of the fields, and of the particles in terms of those fields.

## 6.2 The Low Formalism

Let us start with a general Low-style action, whose kinetic and field terms are those of the Darwin system:

$$\begin{aligned} S[\mathbf{q}, \phi, \mathbf{A}] = \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[ \frac{m\dot{q}^2}{2} + \frac{m\dot{q}^4}{8c^2} - e\phi(\mathbf{q}, t) + \frac{e\dot{\mathbf{q}}}{c} \cdot \mathbf{A}(\mathbf{q}, t) \right] \\ + \frac{1}{8\pi} \int \mathrm{d}t \int \mathrm{d}^3\mathbf{x} \left( E^2 - B^2 \right)_D , \end{aligned} \tag{6.3}$$

where the particle coordinates and velocities are functions of the label $z_0$, thought of as initial conditions. Here $\phi(\mathbf{q}, t)$ and $\mathbf{A}(\mathbf{q}, t)$ is shorthand for writing the fields $\phi(\mathbf{x}, t)$ and $\mathbf{A}(\mathbf{x}, t)$ evaluated at the particle location $\mathbf{q}$. More explicitly, the coupling terms have the form

$$S^{\phi}_{\mathrm{coup}}[\mathbf{q}, \phi, \mathbf{A}] = \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[ -e \int \mathrm{d}^3\mathbf{x} \phi(\mathbf{x}, t) \delta^{(3)} (\mathbf{x} - \mathbf{q}) \right] \tag{6.4}$$

for the electric field term, and similarly for the magnetic coupling. Likewise the subscript $D$ on the field term denotes the Darwin approximated field action as given in (4.43):

$$\begin{aligned}
S_F \; &:= \int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\left(E^2 - B^2\right)_D \\
&= \int \mathrm{d}t\,\mathrm{d}^3\mathbf{x}\left[\frac{2}{c}\frac{\partial\mathbf{A}}{\partial t}\cdot\nabla\phi + \|\nabla\phi\|^2 - \mathbf{A}\cdot\left(\nabla\left(\nabla\cdot\mathbf{A}\right) - \nabla^2\mathbf{A}\right)\right].
\end{aligned} \tag{6.5}$$

In order to familiarize ourselves with variations in this setting, we may compute $\delta S^{\phi}_{\mathrm{coup}}/\delta\mathbf{q}(z_0,t)$. We find

$$\begin{aligned}
\delta S^{\phi}_{\mathrm{coup}} &= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0)\left[-e\int \mathrm{d}^3\mathbf{x}\phi(\mathbf{x},t)\left(\frac{\partial}{\partial\mathbf{q}}\delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\right)\delta\mathbf{q}\right] \\
&= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0)\left[-e\int \mathrm{d}^3\mathbf{x}\phi(\mathbf{x},t)\left(-\frac{\partial}{\partial\mathbf{x}}\delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\right)\delta\mathbf{q}\right] \\
&= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0)\left[-e\int \mathrm{d}^3\mathbf{x}\frac{\partial\phi(\mathbf{x},t)}{\partial\mathbf{x}}\delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\delta\mathbf{q}\right] \\
&= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0)\left[-e\frac{\partial\phi(\mathbf{q},t)}{\partial\mathbf{q}}\right]\delta\mathbf{q},
\end{aligned} \tag{6.6}$$

so that

$$\frac{\delta S^{\phi}_{\mathrm{coup}}}{\delta\mathbf{q}(z_0,t)} = \left(-e\frac{\partial\phi(\mathbf{q},t)}{\partial\mathbf{q}}\right) f_0(z_0). \tag{6.7}$$

Alternately, we may vary with respect to $\phi$ to compute $\delta S^{\phi}_{\mathrm{coup}}/\delta\phi(\bar{\mathbf{x}},t)$.

We have

$$
\begin{aligned}
\delta S^{\phi}_{\text{coup}} &= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[ -e \int \mathrm{d}^3\mathbf{x} \frac{\delta\phi(\mathbf{x},t)}{\delta\phi(\bar{\mathbf{x}},t)} \delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\delta\phi \right] \\
&= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[ -e \int \mathrm{d}^3\mathbf{x}\, \delta^{(3)}\left(\mathbf{x}-\bar{\mathbf{x}}\right) \delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\delta\phi \right] \\
&= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[ -e \delta^{(3)}\left(\bar{\mathbf{x}}-\mathbf{q}\right)\delta\phi \right] \\
&= \int \mathrm{d}t \left[ -e \int \mathrm{d}^6 z_0 f_0(z_0) \delta^{(3)}\left(\bar{\mathbf{x}}-\mathbf{q}\right) \right] \delta\phi. 
\end{aligned} \tag{6.8}
$$

The integral $\mathrm{d}^6 z_0$ can be performed, in principle, since $\mathbf{q} = \mathbf{q}(z_0, t)$. But it is easier to see what this is if we map forward along trajectories. If we look at the map taking $f_0(\mathbf{x}_0, \mathbf{v}_0) \to f(\mathbf{x}, \mathbf{v}, t)$, then we have

$$
\mathbf{x} = \mathbf{q}(z_0, t)\big|_{z_0 = \mathbf{q}^{-1}(\mathbf{x},t)} \tag{6.9}
$$

relating the observation point $\mathbf{x}$ to the trajectory point $\mathbf{q}$. We also use the fact that

$$
\mathrm{d}^6 z_0 = \mathrm{d}^3\mathbf{x}_0\, \mathrm{d}^3\mathbf{v}_0 \to \mathrm{d}^3\mathbf{x}\, \mathrm{d}^3\mathbf{v}. \tag{6.10}
$$

Then the above calculation becomes

$$
\begin{aligned}
\delta S^{\phi}_{\text{coup}} &= \int \mathrm{d}t \left[ -e \int \mathrm{d}^3\mathbf{x}\, \mathrm{d}^3\mathbf{v} f(\mathbf{x},\mathbf{v},t) \delta^{(3)}\left(\bar{\mathbf{x}}-\mathbf{x}\right) \right] \delta\phi \\
&= \int \mathrm{d}t \left[ -e \int \mathrm{d}^3\mathbf{v} f(\bar{\mathbf{x}},\mathbf{v},t) \right] \delta\phi = -\int \mathrm{d}t \rho(\bar{\mathbf{x}},t)\delta\phi,
\end{aligned} \tag{6.11}
$$

which means that

$$
\frac{\delta S^{\phi}_{\text{coup}}}{\delta\phi(\bar{\mathbf{x}},t)} = -\rho(\bar{\mathbf{x}},t). \tag{6.12}
$$

We may now perform the same calculations with the magnetic coupling term

$$S^{\mathbf{A}}_{\text{coup}}[\mathbf{q},\phi,\mathbf{A}] = \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[\frac{e\dot{\mathbf{q}}}{c} \cdot \int \mathrm{d}^3\mathbf{x}\mathbf{A}(\mathbf{x},t)\delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\right]. \quad (6.13)$$

Variation in $\mathbf{q}$ gives

$$\begin{aligned}
\delta S^{\mathbf{A}}_{\text{coup}} &= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \frac{\partial}{\partial q_k}\left(\frac{e\dot{q}_i}{c}\int \mathrm{d}^3\mathbf{x} A_i(\mathbf{x},t)\delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\right)\delta q_k \\
&\quad - \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \frac{d}{dt}\frac{\partial}{\partial \dot{q}_k}\left(\frac{e\dot{q}_i}{c}\int \mathrm{d}^3\mathbf{x} A_i(\mathbf{x},t)\delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\right)\delta q_k \\
&= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[\frac{e\dot{q}_i}{c}\int \mathrm{d}^3\mathbf{x} \frac{\partial A_i(\mathbf{x},t)}{\partial x_k}\delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\right]\delta q_k \\
&\quad - \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[\frac{e}{c}\frac{d}{dt}\int \mathrm{d}^3\mathbf{x} A_k(\mathbf{x},t)\delta^{(3)}\left(\mathbf{x}-\mathbf{q}\right)\right]\delta q_k \\
&= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[\frac{e\dot{q}_i}{c}\frac{\partial A_i(\mathbf{q},t)}{\partial q_k} - \frac{e}{c}\frac{dA_k(\mathbf{q},t)}{dt}\right]\delta q_k, \qquad (6.14)
\end{aligned}$$

where the delta functions have been used to exchange $\partial/\partial q_k$ for $\partial/\partial x_k$, and then the quantities have been integrated by parts. This implies that

$$\frac{\delta S^{\mathbf{A}}_{\text{coup}}}{\delta q_k(z_0,t)} = \left[\frac{e\dot{q}_i}{c}\frac{\partial A_i(\mathbf{q},t)}{\partial q_k} - \frac{e}{c}\frac{dA_k(\mathbf{q},t)}{dt}\right] f_0(z_0). \qquad (6.15)$$

Now as a matter of some interest, we may compute the variation with respect to $\mathbf{A}(\bar{\mathbf{x}}, t)$. Proceeding in this fashion, we have

$$
\begin{aligned}
\delta S^{\mathbf{A}}_{\text{coup}} &= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[ \frac{e\dot{\mathbf{q}}}{c} \cdot \int \mathrm{d}^3\mathbf{x} \frac{\delta \mathbf{A}(\mathbf{x}, t)}{\delta \mathbf{A}(\bar{\mathbf{x}}, t)} \delta^{(3)}\left(\mathbf{x} - \mathbf{q}\right) \delta\mathbf{A} \right] \\
&= \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[ \frac{e\dot{\mathbf{q}}}{c} \cdot \int \mathrm{d}^3\mathbf{x} \delta^{(3)}\left(\mathbf{x} - \bar{\mathbf{x}}\right) \delta^{(3)}\left(\mathbf{x} - \mathbf{q}\right) \delta\mathbf{A} \right] \\
&= \int \mathrm{d}t \left[ \int \mathrm{d}^6 z_0 f_0(z_0) \frac{e\dot{\mathbf{q}}}{c} \delta^{(3)}\left(\bar{\mathbf{x}} - \mathbf{q}\right) \right] \cdot \delta\mathbf{A} \\
&= \int \mathrm{d}t \left[ \int \mathrm{d}^6 z f(z, t) \frac{e\mathbf{v}}{c} \delta^{(3)}\left(\bar{\mathbf{x}} - \mathbf{x}\right) \right] \cdot \delta\mathbf{A},
\end{aligned} \tag{6.16}
$$

where $z = (\mathbf{x}, \mathbf{v})$ and

$$
\mathbf{x} = \mathbf{q}(z, t)\big|_{z=\mathbf{q}^{-1}(\mathbf{x},t)} \qquad \text{and} \qquad \mathbf{v} = \dot{\mathbf{q}}(z, t)\big|_{z=\mathbf{q}^{-1}(\mathbf{x},t)}, \tag{6.17}
$$

with a slight abuse of notation. Then we find

$$
\begin{aligned}
\delta S^{\mathbf{A}}_{\text{coup}} &= \int \mathrm{d}t \left[ \int \mathrm{d}^3\mathbf{x}\, \mathrm{d}^3\mathbf{v} f(\mathbf{x}, \mathbf{v}, t) \frac{e\mathbf{v}}{c} \delta^{(3)}\left(\bar{\mathbf{x}} - \mathbf{x}\right) \right] \cdot \delta\mathbf{A} \\
&= \int \mathrm{d}t \left[ \int \mathrm{d}^3\mathbf{v} f(\bar{\mathbf{x}}, \mathbf{v}, t) \frac{e\mathbf{v}}{c} \right] \cdot \delta\mathbf{A},
\end{aligned} \tag{6.18}
$$

so that

$$
\frac{\delta S^{\mathbf{A}}_{\text{coup}}}{\delta \mathbf{A}(\bar{\mathbf{x}}, t)} = \int \mathrm{d}^3\mathbf{v} f(\bar{\mathbf{x}}, \mathbf{v}, t) \frac{e\mathbf{v}}{c} = \frac{\mathbf{j}(\bar{\mathbf{x}}, t)}{c}. \tag{6.19}
$$

The variations $\delta S_F/\delta\phi$ and $\delta S_F/\delta\mathbf{A}$, combined with (6.12) and (6.19), yield the Darwin field equations.

## 6.3 Low-Darwin Equations of Motion

As in the general calculations, the Low action allows the field action and coupling terms to be combined to get an action that is purely a functional of $\mathbf{q}(z_0, t)$, with an overall factor of $1/2$. The process which follows is: begin with the action (6.3), substitute the solutions of the Darwin field equations in terms of sources, and obtain the action in terms of $\mathbf{q}(z_0, t)$ alone. Though we have seen this process from a general viewpoint, it is nevertheless useful to briefly go through the calculations again in order to see how the distribution function $f$ and the initial conditions $z$ carry through.

For this purpose, we begin by rewriting the field term in the action. Invoking the Coulomb gauge condition from the outset, and dropping the higher order term $(1/c^2)(\partial\mathbf{A}/\partial t)^2$, we have

$$
\begin{aligned}
S_F &= \frac{1}{8\pi}\int \mathrm{d}^3\bar{\mathbf{x}}(E^2 - B^2)_D \\
&= \frac{1}{8\pi}\int \mathrm{d}^3\bar{\mathbf{x}}\left[-\phi_D\nabla^2\phi_D + \mathbf{A}_D\cdot\nabla^2\mathbf{A}_D\right] = \frac{1}{2}\int \mathrm{d}^3\bar{\mathbf{x}}\left[\rho\phi - \frac{\mathbf{j}}{c}\cdot\mathbf{A}\right] \\
&= \frac{1}{2}\int \mathrm{d}^3\bar{\mathbf{x}}\left[\phi_D(\bar{\mathbf{x}}, t)\left(e\int \mathrm{d}^3\mathbf{x}\,\mathrm{d}^3\mathbf{v} f(\mathbf{x}, \mathbf{v}, t)\delta^{(3)}\left(\bar{\mathbf{x}} - \mathbf{x}\right)\right)\right] \\
&\quad -\frac{1}{2c}\int \mathrm{d}^3\bar{\mathbf{x}}\left[\mathbf{A}_D(\bar{\mathbf{x}}, t)\cdot e\int \mathrm{d}^3\mathbf{x}\,\mathrm{d}^3\mathbf{v} f(\mathbf{x}, \mathbf{v}, t)\mathbf{v}\delta^{(3)}\left(\bar{\mathbf{x}} - \mathbf{x}\right)\right].
\end{aligned}
\tag{6.20}
$$

Mapping back to initial conditions, this becomes

$$
\begin{aligned}
S_F &= \frac{1}{2}\int \mathrm{d}^3\bar{\mathbf{x}}\left[\phi_D(\bar{\mathbf{x}}, t)\left(e\int \mathrm{d}^3\mathbf{x}_0\,\mathrm{d}^3\mathbf{v}_0 f_0(\mathbf{x}_0, \mathbf{v}_0)\delta^{(3)}\left(\bar{\mathbf{x}} - \mathbf{q}\right)\right)\right] \\
&\quad -\frac{1}{2c}\int \mathrm{d}^3\bar{\mathbf{x}}\left[\mathbf{A}_D(\bar{\mathbf{x}}, t)\cdot e\int \mathrm{d}^3\mathbf{x}_0\,\mathrm{d}^3\mathbf{v}_0 f_0(\mathbf{x}_0, \mathbf{v}_0)\dot{\mathbf{q}}\delta^{(3)}\left(\bar{\mathbf{x}} - \mathbf{q}\right)\right] \\
&= \frac{1}{2}\int \mathrm{d}^6 z_0 f_0(z_0)\int \mathrm{d}^3\bar{\mathbf{x}}\phi_D(\bar{\mathbf{x}}, t)\delta^{(3)}\left(\bar{\mathbf{x}} - \mathbf{q}\right) \\
&\quad -\frac{1}{2}\int \mathrm{d}^6 z_0 f_0(z_0)\int \mathrm{d}^3\bar{\mathbf{x}}\frac{e\dot{\mathbf{q}}}{c}\cdot\mathbf{A}_D(\bar{\mathbf{x}}, t)\delta^{(3)}\left(\bar{\mathbf{x}} - \mathbf{q}\right).
\end{aligned}
\tag{6.21}
$$

This gives

$$S_F = \frac{1}{2}\int \mathrm{d}^6 z_0 f_0(z_0) \left[e\phi_D(\mathbf{q},t) - \frac{e\dot{\mathbf{q}}}{c}\cdot \mathbf{A}_D(\mathbf{q},t)\right]. \tag{6.22}$$

From this we see explicitly how the distribution function and initial conditions appear and factor out to give the same form as the coupling terms.

We are thus again in a postition to write down a full action, this time for the Low-Darwin system. Using the results of the above calculations to combine the electromagnetic field term with the coupling term, we may write the full, self-consistent action:

$$S_{LD}[\mathbf{q}(z_0,t)] = \int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \left[T(\dot{\mathbf{q}}) + \frac{e}{2}\left(-\phi_D(\mathbf{q}) + \frac{\dot{\mathbf{q}}}{c}\cdot \mathbf{A}_D(\mathbf{q})\right)\right], \tag{6.23}$$

where the function $T$ is for a general kinetic term. In our case this is

$$T(\dot{\mathbf{q}}) = \frac{m\dot{q}^2}{2} + \frac{m\dot{q}^4}{8c^2}, \tag{6.24}$$

but we keep the notation $T$ for later convenience. The arguments of $S_{LD}$ have also been suppressed. Although the action is written in terms of $\phi_D$ and $\mathbf{A}_D$, these are really to be considered shorthand for the solutions of the Darwin field equations in terms of sources. These therefore depend ultimately on the particle coordinates, and so in this sense $S_{LD} = S_{LD}[\mathbf{q}]$ only.

Now Coulomb gauge alone stipulates that $\phi$ solve Poisson's equation.

If we let $K$ denote the kernel of Poisson's equation, then we have

$$
\begin{aligned}
\phi_D(\mathbf{x},t) &= \int \mathrm{d}^3\bar{\mathbf{x}} K\left(\mathbf{x} \mid \bar{\mathbf{x}}\right) e \int \mathrm{d}^3\mathbf{v} f(\bar{\mathbf{x}},\mathbf{v},t) \\
&= \int \mathrm{d}^3\bar{\mathbf{x}} K\left(\mathbf{x} \mid \bar{\mathbf{x}}\right) e \int \mathrm{d}^3\mathbf{x}' \int \mathrm{d}^3\mathbf{v} f(\mathbf{x}',\mathbf{v},t)\delta^{(3)}\left(\bar{\mathbf{x}}-\mathbf{x}'\right) \\
&= \int \mathrm{d}^3\bar{\mathbf{x}} K\left(\mathbf{x} \mid \bar{\mathbf{x}}\right) e \int \mathrm{d}^6 z_0 f_0(z_0)\delta^{(3)}\left(\bar{\mathbf{x}}-\mathbf{q}\right) \\
&= e\int \mathrm{d}^6 z_0 f_0(z_0) K\left(\mathbf{x} \mid \mathbf{q}\right). \qquad (6.25)
\end{aligned}
$$

Then the $\phi$-term in the action is

$$
S^{\phi}_{LD} = -\frac{e}{c}\int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0) \int \mathrm{d}^6 z_0' f_0(z_0') K\left(\mathbf{q} \mid \mathbf{q}'\right). \qquad (6.26)
$$

Looking ahead, we take a moment to obtain the variation of this term with respect to the particle coordinates. We calculate

$$
\begin{aligned}
&\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S^{\phi}_{LD}[\mathbf{q}+\epsilon\delta\mathbf{q}] \\
&= -\frac{e}{2}\int \mathrm{d}^6 z_0 f_0(z_0) \int \mathrm{d}^6 z_0' f_0(z_0') \Bigg[K\left(\mathbf{q} \mid \mathbf{q}'\right) + \epsilon\delta\mathbf{q}\frac{\partial K\left(\mathbf{q} \mid \mathbf{q}'\right)}{\partial \mathbf{q}} \\
&\qquad\qquad + \epsilon\delta\mathbf{q}'\frac{\partial K\left(\mathbf{q} \mid \mathbf{q}'\right)}{\partial \mathbf{q}'} + O(\epsilon^2)\Bigg] \\
&= -\frac{e}{2}\int \mathrm{d}^6 z_0 f_0(z_0) \int \mathrm{d}^6 z_0' f_0(z_0') \left[\frac{\partial K\left(\mathbf{q} \mid \mathbf{q}\right)}{\partial \mathbf{q}} + \frac{\partial \mathbf{q}'}{\partial \mathbf{q}}\frac{\partial K\left(\mathbf{q} \mid \mathbf{q}'\right)}{\partial \mathbf{q}'}\right]\delta\mathbf{q} \\
&= -\frac{e}{2}\int \mathrm{d}^6 z_0 f_0(z_0) \left[2\frac{\partial}{\partial \mathbf{q}}\int \mathrm{d}^6 z_0' f_0(z_0') K\left(\mathbf{q} \mid \mathbf{q}\right)\right]\delta\mathbf{q} \\
&= -e\int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0)\frac{\partial \phi_D(\mathbf{q},t)}{\partial \mathbf{q}}\delta\mathbf{q}, \qquad (6.27)
\end{aligned}
$$

so that

$$\frac{\delta S_{LD}^{\phi}}{\delta \mathbf{q}} = -e\frac{\partial \phi_D(\mathbf{q},t)}{\partial \mathbf{q}} f_0(z_0). \tag{6.28}$$

We may follow a similar procedure for the **A**-term in the action. Since $\mathbf{A}_D$ solves the Darwin field equation, we may express it in terms of the Darwin kernel $K_{ij}^D$ (the subscript $D$ will be written as a superscript to avoid confusion with subscript indices):

$$\begin{aligned}
A_i^D(\mathbf{x},t) &= \int \mathrm{d}^3\bar{\mathbf{x}} K_{ij}^D\left(\mathbf{x} \mid \bar{\mathbf{x}}\right) e \int \mathrm{d}^3\mathbf{v} v_j f(\bar{\mathbf{x}},\mathbf{v},t) \\
&= \int \mathrm{d}^3\bar{\mathbf{x}} K_{ij}^D\left(\mathbf{x} \mid \bar{\mathbf{x}}\right) e \int \mathrm{d}^3\mathbf{x}'\, \mathrm{d}^3\mathbf{v} v_j f(\mathbf{x}',\mathbf{v},t)\delta^{(3)}\left(\bar{\mathbf{x}}-\mathbf{x}'\right) \\
&= \int \mathrm{d}^3\bar{\mathbf{x}} K_{ij}^D\left(\mathbf{x} \mid \bar{\mathbf{x}}\right) e \int \mathrm{d}^6 z_0 f_0(z_0)\dot{q}_j \delta^{(3)}\left(\bar{\mathbf{x}}-\mathbf{x}\right) \\
&= e\int \mathrm{d}^6 z_0 f_0(z_0) K_{ij}^D\left(\mathbf{x} \mid \mathbf{q}\right) \dot{q}_j.
\end{aligned} \tag{6.29}$$

The **A**-term in the action becomes

$$S_{LD}^{\mathbf{A}} = \frac{e}{2c}\int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0)\dot{q}_i \int \mathrm{d}^6 z_0' f_0(z_0') e K_{ij}^D\left(\mathbf{q} \mid \mathbf{q}'\right) \dot{q}_j'. \tag{6.30}$$

Performing the variation over particle coordinates, we compute

$$\left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S^{\mathbf{A}}_{LD}[\mathbf{q}+\epsilon\delta\mathbf{q}]$$

$$= -\frac{e}{2c}\int \mathrm{d}^6 z_0 f_0(z_0) \int \mathrm{d}^6 z_0' f_0(z_0') \Bigg[ 2\delta\dot{q}_i K^D_{ij}\left(\mathbf{q} \mid \mathbf{q}'\right)\dot{q}_j'$$

$$+\delta q_k \frac{\partial K^D_{ij}\left(\mathbf{q} \mid \mathbf{q}'\right)}{\partial q_k}\dot{q}_i\dot{q}_j' + \delta q_l \frac{\partial q_k'}{\partial q_l}\frac{\partial K^D_{ij}\left(\mathbf{q} \mid \mathbf{q}'\right)}{\partial q_k'}\dot{q}_i\dot{q}_j'\Bigg]$$

$$= \frac{e}{c}\int \mathrm{d}t \int \mathrm{d}^6 z_0 f_0(z_0)\left[\frac{\partial}{\partial q_k}\left(q_i A^D_i(\mathbf{q},t)\right) - \frac{dA^D_k(\mathbf{q},t)}{dt}\right]\delta q_k, \qquad (6.31)$$

so that

$$\frac{\delta S^{\mathbf{A}}_{LD}}{\delta q_k} = \frac{e}{c}\left[\frac{\partial}{\partial q_k}\left(q_i A^D_i(\mathbf{q},t)\right) - \frac{dA^D_k(\mathbf{q},t)}{dt}\right] f_0(z_0). \qquad (6.32)$$

All that remains to compute the equations of motion is variation of the kinetic term $T(\dot{\mathbf{q}})$. Since this only depends on $\dot{\mathbf{q}}$, we find

$$\left(\frac{\partial}{\partial \mathbf{q}} - \frac{d}{dt}\frac{\partial}{\partial \dot{\mathbf{q}}}\right) T(\dot{\mathbf{q}}) = -\frac{d}{dt}T'(\dot{\mathbf{q}}). \qquad (6.33)$$

Putting together the various variations, we then have the equations of motion for the Low-Darwin system:

$$-\frac{d}{dt}T_k'(\dot{\mathbf{q}}) - e\frac{\partial \phi_D(\mathbf{q},t)}{\partial q_k} + \frac{e}{c}\frac{\partial}{\partial q_k}\left(\dot{q}_i A^D_i(\mathbf{q},t)\right) - \frac{e}{c}\frac{dA^D_k(\mathbf{q},t)}{dt} = 0. \qquad (6.34)$$

Using the vector identity

$$\nabla(\mathbf{A}\cdot\mathbf{B}) = \mathbf{A}\times(\nabla\times\mathbf{B}) + \mathbf{B}\times(\nabla\times\mathbf{A}) + (\mathbf{A}\cdot\nabla)\mathbf{B} + (\mathbf{B}\cdot\nabla)\mathbf{A}, \quad (6.35)$$

we may write

$$\nabla(\dot{\mathbf{q}}\cdot\mathbf{A}^D) = \dot{\mathbf{q}}\times\mathbf{B}^D + (\dot{\mathbf{q}}\cdot\nabla)\mathbf{A}^D. \qquad (6.36)$$

The last term combines with the $d\mathbf{A}^D/dt$ in the equation of motion to leave $\partial\mathbf{A}^D/\partial t$. Introducing the notation $\mathbf{E}^T := -(1/c)(\partial\mathbf{A}^D/\partial t)$ and $\mathbf{E}^L := -\nabla\phi^D$, we may finally bring the equation of motion to the form

$$\frac{d}{dt}T'(\dot{\mathbf{q}}) = e\left[\mathbf{E}^L + \mathbf{E}^T + \frac{\dot{\mathbf{q}}}{c}\times\mathbf{B}^D\right]. \tag{6.37}$$

## 6.4 Low-Darwin Kinetic Term

Up to this point, the kinetic term $T(\dot{\mathbf{q}})$ has not played a crucial role in the calculations. It has merely come along for the ride. In this section we investigate the implications of using the usual Darwin kinetic term. Several papers in the literature simply use $m\dot{q}^2/2$ as the kinetic term for the Darwin field system. From our calculations with the particle system, this is certainly not to be expected. We should therefore check what impact the additional $m\dot{q}^4/8c^2$ term has on the equations of motion. Given the Darwin kinetic term

$$T(\dot{\mathbf{q}}) = \frac{m\dot{q}^2}{2} + \frac{m\dot{q}^4}{8c^2}, \tag{6.38}$$

it is straighforward to compute

$$T'(\dot{\mathbf{q}}) = \left(1 + \frac{\dot{q}^2}{2c^2}\right)m\dot{\mathbf{q}}. \tag{6.39}$$

Computing the time-derivative

$$\begin{aligned}\frac{dT'_k}{dt} &= m\ddot{q}_i\left[\delta_{ik}\left(1 + \frac{\dot{q}^2}{2c^2}\right) + \frac{\dot{q}_i\dot{q}_k}{c^2}\right]\\ &= m\ddot{q}\left[\delta_{ij} + a_{ij}\left(\frac{\dot{q}^2}{c^2}\right)\right] =: m\ddot{q}A_{ij},\end{aligned} \tag{6.40}$$

where $a_{ij} := \delta_{ij}/2 + x_i x_j$ with $\mathbf{x} := \dot{\mathbf{q}}/\left\|\dot{\mathbf{q}}\right\|$. Thus $a_{ij}$ is of order unity. Our task is to find the inverse of the matrix $A_{ij}$.

We need only work to second order in $v/c$. We therefore write

$$A_{ij}^{-1} =: B_{ij} := \delta_{ij} + b_{ij}\left(\frac{\dot{q}^2}{c^2}\right). \tag{6.41}$$

By definition of a matrix inverse, we have

$$\begin{aligned}\delta_{ik} = A_{ij}B_{jk} &= \left[\delta_{ij} + a_{ij}\left(\frac{\dot{q}^2}{c^2}\right)\right]\left[\delta_{ik} + b_{ik}\left(\frac{\dot{q}^2}{c^2}\right)\right] \\ &= \delta_{ik} + a_{ik}\left(\frac{\dot{q}^2}{c^2}\right) + b_{ik}\left(\frac{\dot{q}^2}{c^2}\right) + O\left(\frac{\dot{q}^4}{c^4}\right)\end{aligned} \tag{6.42}$$

to the order we are considering. Thus $b_{ik} = -a_{ik}$ and the inverse is $A_{ij}^{-1} = \delta_{ij} - a_{ij}(\dot{q}^2/c^2)$. Using this, we may write the Low-Darwin equation of motion as

$$\begin{aligned}m\ddot{q}_i &= \left[\delta_{ij} - a_{ij}\left(\frac{\dot{q}^2}{c^2}\right)\right]\left[eE_i^L + eE_j^T + \left(\frac{e\dot{\mathbf{q}}}{c} \times \mathbf{B}^D\right)_j\right] \\ &= \left(\delta_{ij} - a_{ij}\left(\frac{\dot{q}^2}{c^2}\right)\right)E_j^L + eE_i^T + \frac{e}{c}\left[\dot{\mathbf{q}} \times \mathbf{B}^D\right]_i,\end{aligned} \tag{6.43}$$

where terms of higher order have been dropped.[1] Writing $\mathbf{E} = \mathbf{E}^L + \mathbf{E}^T$, this gives

$$m\ddot{q}_i = eE_i + \frac{e}{c}\left[\dot{\mathbf{q}} \times \mathbf{B}\right]_i - e\left(\frac{\delta_{ij}}{2} + x_i x_j\right)\left(\frac{\dot{q}^2}{c^2}\right)E_j^L. \tag{6.44}$$

[1]Recall that $\mathbf{E}^T := -(1/c)(\partial\mathbf{A}^D/\partial t)$. Since $\mathbf{A}^D \sim 1/c$, then $\mathbf{E}^T$ is already of order $1/c^2$. Contracting this with $A_{ij}^{-1}$ and keeping only terms up to order-$(v/c)^2$ leaves only the term $\delta_{ij}E_j^T = E_i^T$.

There is no obvious reason why the last term should be zero to this order, so the use of only $m\dot{q}^2/2$ for the Darwin kinetic term is not consistent, and the corrective $m\dot{q}^4/8c^2$ term must be retained.

# Chapter 7

# Conclusions

The treatment of the D2BP has resulted in a few noteworthy observations. One glaring result is the incompatibility of the methods of symplectic reduction as applied to the N2BP as opposed to those same methods applied to the D2BP. We saw that the value $\alpha$ determining the level sets of the momentum map for $(\mathbb{R}^3, +)$ may be carried along without impinging upon the second reduction under $SO(3)$ in the case of the N2BP, but that this value has a definite impact on the second reduction in the case of the D2BP. There are two reasonable tracks to follow for a resolution. In one manner of reasoning, we may act as if the first reduction never occurred. In this scenario, we are merely given a Hamiltonian which we know nothing about, save that $\alpha$ plays the role of three scalar parameters, on the same level as the mass and charge parameters. Given this state of affairs, this Hamiltonian is in fact *not* invariant under the diagonal $SO(3)$ action, unless $\alpha \equiv \mathbf{0}$. Backtracking, and only *then* realizing that this Hamiltonian was indeed the result of a reduction under $(\mathbb{R}^3, +)$, we find that the case $\alpha \equiv \mathbf{0}$ in fact corresponds to the choice of moving to the center-of-mass frame. From this point of view, the center-of-mass frame is singled out as preferred, contrary to what was found in the study of the N2BP.

On the other hand, we may allow ourselves knowledge of reduction at each stage. We are aware of the reduction under $(\mathbb{R}^3, +)$, cognizant of the fact that, under this reduction, the moment map associates the total momentum $\mathbf{p}$ with elements of the dual to the Lie algebra of $(\mathbb{R}^3, +)$, and that $\alpha$ lies in this space as well. Then the subsequent action of $SO(3)$, when acting on $\mathbf{p}$, must also act on $\alpha$. In this case, then, the Hamiltonian *is* invariant under $SO(3)$, even when $\alpha \neq \mathbf{0}$. From this point of view, the center-of-mass frame is not singled out as preferential.

The resolution of the two points of view seems to come in the analysis of the interaction. Since the interaction of the D2BP is momentum-dependent, and since momentum is frame-dependent, then the choice of reference frame could very well impact the symmetry of the system. This is not the case with the N2BP, since the interaction is only dependent on relative separation, not on momentum. From this perspective, the terms with $\|(\text{const})\alpha - \mathbf{p}\|^4$ are ambiguous. Their origin in the derivation comes from what is essentially a kinetic term; however their effect in reference to the preceding argument is one of an interaction term.

We also found that the D2BP allows for the mathematical possibility of three fixed points for the equations of motion, leading to three circular orbits. The subsequent estimates, however, showed that the conditions for the existence of all three roots were incongruent with the requirement that the Darwin system be a good approximation to the relativistic system. The requirement that $(v/c) \ll 1$ precludes the existence of two of the three real roots. Thus the Darwin system, taken in and of itself as a self-standing system, permits three circular orbits given high enough velocities. But taken as a valid approximation to a relativistic system, the conditions on the velocities forbid two of the circular orbits and we are left with only one.

The remainder of the work focused on what was here termed the self-consistent Darwin system. The original derivation of the standard Darwin system takes the fields $\phi$ and $\mathbf{A}$ as given. The fact that they satisfy Maxwell's equations is imposed from outside. The self-consistent approach attempts to relieve this situation by starting with an action which includes particles and fields in an inclusive whole. The idea is a simple one. It is a top-down formulation of what was originally a bottom-up procedure. Darwin's original derivation of the system came about by considering a single charge in an external field, then another charge added to this, then another charge, and so on. In this way, Darwin constructed his system particle by particle. The approach presented here, by contrast, attackes the problem from the opposite direction. Here we assume given the action for particles coupled to the electromagnetic field. The equations for the fields are not 'given' per se, but are derived from the action by variation on the fields themselves. Thus in our action, the fields themselves are initially as dynamic as the particles. Given this starting point, we then study the sytem in the situation where the fields are the fields of the particles themselves, and so we may express the potentials in terms of the particle variables. This procedure of particlization leads us eventually to an action which is solely in terms of particles:

$$S_D[\mathbf{q}_a] = \sum_a \int \mathrm{d}t \left( \frac{m_a \dot{q}_a^2}{2} + \frac{m_a \dot{q}_a^4}{8c^2} \right) - \frac{1}{2} \sum_{\substack{a,b \\ b \neq a}} \int \mathrm{d}t \frac{e_a e_b}{\|\mathbf{q}_a - \mathbf{q}_b\|}$$
$$+ \frac{1}{2} \sum_{\substack{a,b \\ b \neq a}} \int \mathrm{d}t \frac{e_a e_b}{2c^2 \|\mathbf{q}_a - \mathbf{q}_b\|} \left[ \dot{\mathbf{q}}_a \cdot \dot{\mathbf{q}}_b + (\dot{\mathbf{q}}_a \cdot \hat{\mathbf{r}}_{ab}) (\dot{\mathbf{q}}_b \cdot \hat{\mathbf{r}}_{ab}) \right]. \tag{7.1}$$

This agrees with the action found elsewhere in the literature.

Another novel feature of the present work is the search for a direct method of establishing the equivalence of the Darwin field equations and the quasistatic solutions. That is, Jackson shows a method of arriving at the form of the vector potential which arises in Darwin's Lagrangian. The first step is to obtain the retarded potential in Lorenz gauge, then transform this solution to Coulomb gauge, and then finally to take the instantaneous, or 'quasistatic', limit of this solution. This yields $\mathbf{A}_C^{qs}$. Other researchers, by contrast, argue the form of the equations before obtaining a solution. The line of resoning is to start with the equations for the potentials in Coulomb gauge, then argue the instantaneous form of these — *i.e.* dropping the term $(1/c^2)(\partial^2 \mathbf{A}/\partial t^2)$ — and solve the resulting equation. This results in $\mathbf{A}_D$. Though the logic of each argument is headed in the same direction, as equations go, it is not altogether certain the the solutions must be the same (though of course it is hoped). In this work we have showed directly the equivalence of $\mathbf{A}_C^{qs}$ and $\mathbf{A}_D$.

The calculation which establishes this equivalence, however, relies on one integral, which has here been termed the Darwin integral. For all practical purposes, this integral is divergent. But there is a consistent way to extract a finite contribution from this integral, and it turns out that this finite contribution is exactly what is needed to establish the equivalence of the two forms of the vector potential. The integral can be calculated in two fashions, one involving the transformation of the problem to a Gauss's law calculation. This method illustrates that the dropping of the infinite contribution is analogous to choosing the zero of a potential at zero rather than infinity. The integral is itself an interesting calculation, but it has no real merit for that. Its main importance is in giving clues to what is lost when truncating an electromagnetic system to Darwin form, and to the difference between the two methods of reching the form of the potential, one giving $\mathbf{A}_C^{qs}$, the other giving $\mathbf{A}_D$.

Finally the Low-Darwin system was an attempt to construct a self-consistent Darwin theory coupled to the Vlasov equation. This follows through straighforwardly. We noted at the time that several papers merely write the force law as $d\mathbf{v}/dt = e\mathbf{E}^L + e\mathbf{E}^T + (e/c)\mathbf{v} \times \mathbf{B}^D$. As was seen, this formulation is not entirely consistent with the Darwin approximation, omitting a term depending on $(v/c)^2\mathbf{E}^L$. Removal of this term requires further arguments.

# Appendices

# Appendix A

# Groundwork for the Two-Body Problem

Before we actually work out the reduction for the two-body problem, it will be helpful to store up some of the necessary concepts at the outset. Since the two-body problem uses reduction via the groups $(\mathbb{R}^3, +)$ and $SO(3)$, we should gather all the information we need about those groups.

## A.1 $(\mathbb{R}^3, +)$ — The Group of Translations

We may write elements of $(\mathbb{R}^3, +)$ as $\mathbf{t} := (t_1, t_2, t_3)$. Addition is performed as usual:

$$(t_1, t_2, t_3) + (s_1, s_2, s_3) = (t_1 + s_1, t_2 + s_2, t_3 + s_3) \ . \tag{A.1}$$

Of course $(0, 0, 0)$ is the identity element.

It will be useful to note that $(\mathbb{R}^3, +)$ is homomorphic to the group of $3 \times 3$ diagonal matrices with strictly positive entries, endowed with matrix multiplication, under the map

$$(t_1, t_2, t_3) \longleftrightarrow \begin{pmatrix} e^{t_1} & 0 & 0 \\ 0 & e^{t_2} & 0 \\ 0 & 0 & e^{t_3} \end{pmatrix} \ . \tag{A.2}$$

## A.2 Lie Algebra of $(\mathbb{R}^3, +)$

Using the correspondence

$$(\mathbb{R}^3, +) \cong \left\{ \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & a_3 \end{pmatrix} \middle| \, a_1, a_2, a_3 \in \mathbb{R}^+ \right\} , \tag{A.3}$$

with matrix multiplication, we may proceed to find the Lie algebra of $(\mathbb{R}^3, +)$. (Matrix entries left blank are zero.)

**Claim 3.** *The Lie algebra of* $(\mathbb{R}^3, +)$ *can be viewed as the set*

$$\left\{ \begin{pmatrix} \xi_1 & & \\ & \xi_2 & \\ & & \xi_3 \end{pmatrix} \middle| \, \xi_1, \xi_2, \xi_3 \in \mathbb{R} \right\} , \tag{A.4}$$

*with* $[\cdot \ , \ \cdot]$ *the usual matrix commutator.*

*Proof :* To see that the Lie algebra contains this set, take a path

$$T(s) = \begin{pmatrix} e^{as} & & \\ & e^{bs} & \\ & & e^{cs} \end{pmatrix} \subseteq (\mathbb{R}^3, +) \ . \tag{A.5}$$

Differentiating at the identity, we have

$$T'(0) = \begin{pmatrix} a & & \\ & b & \\ & & c \end{pmatrix} \ , \tag{A.6}$$

where $a, b, c$ may have any real values.

Going the other direction, suppose

$$T'(0) = \begin{pmatrix} \xi_1 & & \\ & \xi_2 & \\ & & \xi_3 \end{pmatrix} \ . \tag{A.7}$$

Then exponentiating, we have

$$T(s) = \exp\left(T'(0)s\right) = \begin{pmatrix} e^{\xi_1 s} & & \\ & e^{\xi_2 s} & \\ & & e^{\xi_3 s} \end{pmatrix} . \tag{A.8}$$

For any matrix group, the Lie algebra bracket is the matrix commutator $[\cdot\ ,\ \cdot]$. $\square$

## A.3 One-Parameter Subgroup of $(\mathbb{R}^3, +)$

Working in the matrix representation, we may write a one-parameter subgroup of $(\mathbb{R}^3, +)$ as

$$G_\xi = \left\{ e^{s\xi} \middle| \, s \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} e^{s\xi_1} & & \\ & e^{s\xi_2} & \\ & & e^{s\xi_3} \end{pmatrix} \middle| \, s \in \mathbb{R} \right\} . \tag{A.9}$$

In the vector form of $(\mathbb{R}^3, +)$ this is equivalent to

$$G_\xi = \{ (s\xi_1, s\xi_2, s\xi_3) | \, s \in \mathbb{R} \} = \{ s\xi \, | \, s \in \mathbb{R} \} . \tag{A.10}$$

## A.4 Inner Product on the Lie Algebra of $(\mathbb{R}^3, +)$

The Lie algebra of $(\mathbb{R}^3, +)$ has a natural inner product structure using the usual dot product. In terms of the matrices, this involves the trace:

$$(\xi \mid \eta) = \mathrm{t}r \left[ \begin{pmatrix} \xi_1 & & \\ & \xi_2 & \\ & & \xi_3 \end{pmatrix} \cdot \begin{pmatrix} \eta_1 & & \\ & \eta_2 & \\ & & \eta_3 \end{pmatrix} \right] = \xi_1\eta_1 + \xi_2\eta_2 + \xi_3\eta_3 . \tag{A.11}$$

This gives a natural way of switching to the dual of the Lie algebra.

## A.5 Moment Map for $(\mathbb{R}^3,+)$-Action on $\mathbb{R}^6$

Take $(\mathbb{R}^3,+)$ as the group of translations on $(M,\omega)$, where

$$M = \{(\mathbf{p},\mathbf{r})\} = \mathbb{R}^6 \qquad \text{and} \qquad \omega = \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r} \ . \tag{A.12}$$

The moment map $\Phi$ is a map

$$\Phi : M \rightarrow \left(\mathbb{R}^3\right)^* \ , \qquad \text{that is} \qquad \Phi : \mathbb{R}^6 \rightarrow \mathbb{R}^3 \ , \tag{A.13}$$

satisfying

$$\omega(\xi_M,\cdot) = -\,\mathrm{d}\Phi^\xi \qquad \text{and} \qquad \xi_M f(m) = \left.\frac{d}{dt}\right|_{t=0} f\left(e^{t\xi}m\right) \tag{A.14}$$

for all $\xi \in \mathfrak{g} = \mathbb{R}^3$.

The $(\mathbb{R}^3,+)$-action on $M$ is given by

$$\begin{aligned} S_g : M &\rightarrow M \\ (\mathbf{p},\mathbf{r}) &\mapsto (\mathbf{p},\mathbf{r}+g) \ , \end{aligned} \tag{A.15}$$

where $g \in (\mathbb{R}^3,+)$. In this case we write $g = t\xi$ by virtue of (A.10). Using this we calculate for the second of the conditions (A.14)

$$\begin{aligned} \xi_M f(\mathbf{p},\mathbf{r}) &= \left.\frac{d}{dt}\right|_{t=0} f(\mathbf{p},\mathbf{r}+t\xi) \\ &= \frac{\partial f}{\partial r_1}\xi_1 + \frac{\partial f}{\partial r_2}\xi_2 + \frac{\partial f}{\partial r_3}\xi_3 = \xi \cdot \frac{\partial f}{\partial \mathbf{r}} \ . \end{aligned} \tag{A.16}$$

Hence

$$\xi_M = \xi \cdot \frac{\partial}{\partial \mathbf{r}} \ . \tag{A.17}$$

Now look at the first condition in (A.14). Using the above result, the left-hand side becomes

$$\omega(\xi_M,\cdot) = \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r}(\xi_M,\cdot) = \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r}\left(\xi \cdot \frac{\partial}{\partial \mathbf{r}},\cdot\right)$$

$$= -\xi \cdot \mathrm{d}\mathbf{p}(\cdot) = -\mathrm{d}(\xi \cdot \mathbf{p})(\cdot)\ . \tag{A.18}$$

Comparing this with the right-hand side of the same condition in (A.14), we find

$$\Phi^\xi(\mathbf{p},\mathbf{r}) = \mathbf{p}\cdot\xi = \langle\ \mathbf{p}\ ,\ \xi\ \rangle\ . \tag{A.19}$$

Comparing this, in turn, with (2.37), we may write

$$\Phi^\cdot(\mathbf{p},\mathbf{r}) = \langle\ \mathbf{p}\ ,\ \cdot\ \rangle = \mathbf{p}\ , \tag{A.20}$$

where the last equality is somewhat of an abuse of notation. An even greater abuse of notation is the form in which this is most commonly written:

$$\Phi : (\mathbf{p},\mathbf{r}) \mapsto \mathbf{p} \in \mathfrak{g}^* = \mathbb{R}^3\ , \tag{A.21}$$

so written (1) because it is allowed by the isomorphism $\left[\text{Lie alg. of } (\mathbb{R}^3,+)\right]^* \cong \mathbb{R}^3$, and (2) because $\mathbf{p}$ is the physical quantity of interest. With this definition, then

$$\Phi^\xi(\mathbf{p},\mathbf{r}) = \langle\ \mathbf{p}\ ,\ \xi\ \rangle = \mathbf{p}\cdot\xi \in \mathbb{R}\ . \tag{A.22}$$

Hence the moment map for the $(\mathbb{R}^3,+)$ translation action on $\{(\mathbf{p},\mathbf{r})\}$ is

$$\begin{aligned} \Phi : M &\to \mathfrak{g}^* \\ (\mathbf{p},\mathbf{r}) &\mapsto \mathbf{p}\ . \end{aligned} \tag{A.23}$$

## A.6 $SO(3)$ and $\mathfrak{so}(3)$ — The Group of Rotations and its Lie Algebra

The *special orthogonal group*, or *group of rotations*, in three dimensions is the group

$$\begin{aligned} SO(3) &:= \left\{g \in Gl(3,\mathbb{R}) \mid g^T g = \mathrm{id} \qquad \text{and} \qquad \det g = 1\right\} \\ &= \{3 \times 3 \text{ rotation matrices with determinant } 1\} \\ &= \{3 \times 3 \text{ matrices} \mid \text{columns form an orthonormal basis} \\ &\qquad\qquad \text{obeying the right-hand rule}\}\ . \end{aligned} \tag{A.24}$$

The Lie algebra of $SO(3)$ is the set of *infinitesimal rotations.* To figure out the form of these matrices, we have the following.

**Claim 4.** $\mathfrak{so}(3)$ *is given by*

$$\begin{aligned} \mathfrak{so}(3) &= \{\xi \mid \xi \textit{ is a } 3 \times 3 \textit{ skew-symmetric real matrix}\} \\ &= \left\{ \begin{pmatrix} 0 & \xi_z & -\xi_y \\ -\xi_z & 0 & \xi_x \\ \xi_y & -\xi_x & 0 \end{pmatrix} \middle| \, \xi_x, \xi_y, \xi_z \in \mathbb{R} \right\}\ . \end{aligned} \tag{A.25}$$

*Proof :* To prove $\mathfrak{so}(3)$ contains the skew-symmetric matrices, suppose

$$\begin{aligned} &R : \mathbb{R} \to SO(3)\ , \\ &R(0) = \mathrm{id} \end{aligned} \tag{A.26}$$

is a path through the idintity in $SO(3)$. Since each matrix $R(s)$ is in $SO(3)$, we have

$$R(s)^T R(s) = \text{id}. \tag{A.27}$$

Differentiating both sides at $s = 0$, and noting $R(0) = \text{id} = R(0)^T$, we have

$$\begin{aligned} R'(0)^T R(0) + R(0)^T R'(0) &= 0 \; ; \\ R'(0)^T + R'(0) &= 0 \; ; \\ R'(0)^T &= -R'(0), \end{aligned} \tag{A.28}$$

which implies $R'(0) \in \mathfrak{so}(3)$ is skew-symmetric.

To prove the other direction, let $\xi$ be any $3 \times 3$ skew-symmetric matrix. We may define

$$\begin{aligned} R : \mathbb{R} &\to SO(3) \; ; \\ R(s) &= e^{s\xi} \; . \end{aligned} \tag{A.29}$$

To see this is in fact contained in $SO(3)$, we must check first the transpose condition:

$$\left(e^{s\xi}\right)^T \left(e^{s\xi}\right) = e^{s(\xi)^T} e^{s\xi} = e^{-s\xi} e^{s\xi} = \text{id} \; . \tag{A.30}$$

Also the determinant condition must be checked. Using this last result, we find

$$\det\left(e^{s\xi}\right)^T \det\left(e^{s\xi}\right) = \det(\mathrm{id}) \ ;$$

$$\left[\det\left(e^{s\xi}\right)\right]^2 = 1 \ ;$$

$$\det\left(e^{s\xi}\right) = \pm 1. \tag{A.31}$$

Since $R(0) = \mathrm{id}$, then $\det R(0) = 1$. Since $R(s)$ is continuous, and det is a continuous function, then $\det(R(s))$ is continuous. Hence $\det(R(s))$ cannot change from 1 to $-1$. Thus

$$\det\left(e^{s\xi}\right) = 1. \tag{A.32}$$

We must also check that

$$\left.\frac{d}{ds}\right|_{s=0} R(s) = \xi \, . \tag{A.33}$$

We have

$$\left.\frac{d}{ds}\right|_{s=0} R(s) = \left.\frac{d}{ds}\right|_{s=0} e^{s\xi} = \xi e^{s\xi}\big|_{s=0} = \xi \, . \tag{A.34}$$

Hence the two sets are equal. □

Again, since this is a matrix group, the Lie bracket is the usual matrix commutator $[\cdot \ , \cdot]$. Using the identification with $\mathbb{R}^3$ given by

$$\begin{pmatrix} 0 & \xi_z & -\xi_y \\ -\xi_z & 0 & \xi_x \\ \xi_y & -\xi_x & 0 \end{pmatrix} \longleftrightarrow \begin{pmatrix} \xi_x \\ \xi_y \\ \xi_z \end{pmatrix} \ , \tag{A.35}$$

we find

$$[\xi \ , \eta] \longleftrightarrow \hat{\xi} \times \hat{\eta} \ , \tag{A.36}$$

with $\times$ the usual cross product on $\mathbb{R}^3$. Thus

$$(\mathfrak{so}(3), [\ ,\ ]) \cong (\mathbb{R}^3, \times)\,. \tag{A.37}$$

Another useful identity is

$$\xi\hat{\xi} = \begin{pmatrix} 0 & \xi_z & -\xi_y \\ -\xi_z & 0 & \xi_x \\ \xi_y & -\xi_x & 0 \end{pmatrix} \begin{pmatrix} \xi_x \\ \xi_y \\ \xi_z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\ . \tag{A.38}$$

## A.7 One-Parameter Subgroups of $SO(3)$

The one-parameter subgroups of $SO(3)$ have the form

$$\tilde{G}_\xi = \left\{ e^{s\xi} \middle|\ s \in \mathbb{R},\ \xi \in \mathfrak{so}(3) \right\}\ . \tag{A.39}$$

Of particular interest are the subgroups

$$\begin{aligned}
\tilde{G}_{\xi_z} &= \left\{ \exp\left[ s \begin{pmatrix} 0 & \xi_z & 0 \\ -\xi_z & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] \middle|\ s \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \cos(s) & \sin(s) & 0 \\ -\sin(s) & \cos(s) & 0 \\ 0 & 0 & 1 \end{pmatrix} \middle|\ s \in \mathbb{R} \right\}\ , \\
\tilde{G}_{\xi_y} &= \left\{ \exp\left[ s \begin{pmatrix} 0 & 0 & -\xi_y \\ 0 & 0 & 0 \\ \xi_y & 0 & 0 \end{pmatrix} \right] \middle|\ s \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \cos(s) & 0 & -\sin(s) \\ 0 & 1 & 0 \\ \sin(s) & 0 & \cos(s) \end{pmatrix} \middle|\ s \in \mathbb{R} \right\}\ , \\
\tilde{G}_{\xi_z} &= \left\{ \exp\left[ s \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \xi_x \\ 0 & -\xi_x & 0 \end{pmatrix} \right] \middle|\ s \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(s) & \sin(s) \\ 0 & -\sin(s) & \cos(s) \end{pmatrix} \middle|\ s \in \mathbb{R} \right\}\ ,
\end{aligned} \tag{A.40}$$

rotations about the $z$, $y$ and $x$ axes respectively.

## A.8 Inner Product on $\mathfrak{so}(3)$

The Lie algebra of $SO(3)$ may be given an inner product structure by defining

$$(\xi \mid \eta) := -\frac{1}{2}\mathrm{t}r\,[\xi\eta] \ . \tag{A.41}$$

Writing this out, we find

$$\begin{aligned}(\xi \mid \eta) &= -\frac{1}{2}\mathrm{t}r\left[\begin{pmatrix} -(\xi_z\eta_z+\xi_y\eta_y) & -\xi_y\eta_x & \xi_z\eta_x \\ \xi_z\eta_y & -(\xi_z\eta_z+\xi_y\eta_y) & \xi_z\eta_y \\ \xi_x\eta_z & \xi_y\eta_z & -(\xi_z\eta_z+\xi_y\eta_y) \end{pmatrix}\right] \\ &= \xi_x\eta_x+\xi_y\eta_y+\xi_z\eta_z = \hat{\xi}\cdot\hat{\eta} \ . \end{aligned} \tag{A.42}$$

Hence

$$\left(\mathfrak{so}(3), -\frac{1}{2}\mathrm{trace}\right) \cong (\mathbb{R}^3, \cdot) \ . \tag{A.43}$$

Note that with this inner product we may make the identification

$$\mathfrak{so}(3)^* \cong \mathfrak{so}(3), \tag{A.44}$$

of $\mathfrak{so}(3)$ with its dual space, by writing

$$\xi^* := (\xi \mid \cdot) = -\frac{1}{2}\mathrm{t}r\,[\xi\cdot] = \hat{\xi}\cdot(\ \ ) \ . \tag{A.45}$$

$\xi^*$ is the element of $\mathfrak{so}(3)^*$ dual to $\xi \in \mathfrak{so}(3)$.

## A.9 Moment Map for $SO(3)$-Action on $\mathbb{R}^6$

We now take $SO(3)$ as the group of rotations on $(M,\omega)$, where $M = \{(\mathbf{p},\mathbf{r})\} = \mathbb{R}^6$ and $\omega = \mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r}$. Recall that we use the *diagonal action* given by

$$S_g : M \to M \ ;$$

$$(\mathbf{p}, \mathbf{r}) \mapsto (g\mathbf{p}, g\mathbf{r}) \ . \tag{A.46}$$

The terminology "diagonal action" arises from writing the action in block matrix form with copies of $g$ on the diagonal:

$$\begin{pmatrix} \mathbf{p} \\ \mathbf{r} \end{pmatrix} \mapsto \left(\begin{array}{c|c} g & \\ \hline & g \end{array}\right) \begin{pmatrix} \mathbf{p} \\ \mathbf{r} \end{pmatrix} \ . \tag{A.47}$$

Now "building" the moment map as done in the case of $(\mathbb{R}^3, +)$ is a little difficult here. Just going through the equations is a little daunting. Rather, we should try to use a little reason to make an educated guess at the form of the momentum map, then check to see if it works out.

First, notice that $\Phi : M {\to} \mathfrak{so}(3)^*$. By our trace formula for the inner product on $\mathfrak{so}(3)$, we can see that $\mathfrak{so}(3) \cong \mathfrak{so}(3)^*$ — the inner product also gives us a pairing. This means

$$\Phi : M \to \mathfrak{so}(3)^* \cong \mathfrak{so}(3) \ ;$$

$$(\mathbf{p}, \mathbf{r}) \mapsto \text{some skew-symmetric matrix.} \tag{A.48}$$

Now this skew-symmetrc matrix in $\mathfrak{so}(3)^*$ must involve only the components of $\mathbf{p}$ and $\mathbf{r}$, since that is the only data we have to put in. So how do we make a skew-symmetric matrix from $(\mathbf{p}, \mathbf{r})$?

First, how do we make a matrix at all? Well, we could write

$$\mathbf{p}\mathbf{r}^T = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} \begin{pmatrix} r_x \; r_y \; r_z \end{pmatrix} = \begin{pmatrix} p_x r_x & p_x r_y & p_x r_z \\ p_y r_x & p_y r_y & p_y r_z \\ p_z r_x & p_z r_y & p_z r_z \end{pmatrix} . \tag{A.49}$$

But this is hardly skew-symmetric. Nor is $\mathbf{r}\mathbf{p}^T$. But the combination

$$\mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T = \begin{pmatrix} 0 & p_x r_y - r_x p_y & p_x r_z - r_x p_z \\ p_y r_x - r_y p_x & 0 & p_y r_z - r_y p_z \\ p_z r_x - r_z p_x & p_z r_y - r_z p_y & 0 \end{pmatrix} \tag{A.50}$$

*is* skew-symmetric! With this in mind, we will try the following: define $\Phi$ by

$$\Phi : M \longrightarrow \mathfrak{so}(3)^* \; ;$$

$$(\mathbf{p}, \mathbf{r}) \mapsto (\mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T) \, . \tag{A.51}$$

In this case the function $\Phi^{\cdot}(\mathbf{p}, \mathbf{r}) : \mathfrak{so}(3)^* \!\rightarrow\! \mathbb{R}$ is given by

$$\Phi^\xi(\mathbf{p}, \mathbf{r}) = \langle \, (\mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T) \, , \, \xi \, \rangle = \big( (\mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T) \; \big| \; \xi \big) = -\frac{1}{2} \mathrm{t}r \left[ \xi (\mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T) \right] . \tag{A.52}$$

We may also express this as

$$\begin{aligned} \Phi^\xi(\mathbf{p}, \mathbf{r}) &= -\frac{1}{2} \mathrm{t}r \left[ \xi (\mathbf{p}\mathbf{r}^T - \mathbf{r}\mathbf{p}^T) \right] \\ &= - \left[ \xi_x (r_y p_z - r_z p_y) + \xi_y (r_z p_x - r_x p_z) + \xi_z (r_x p_y - r_y p_x) \right] \\ &= -\hat{\xi} \cdot (\mathbf{r} \times \mathbf{p}). \end{aligned} \tag{A.53}$$

Now let us check to see that the properties of a moment map are satisfied.

First let us calculate the vector $\xi_M$ from the requirement (2.41):

$$\xi_M f(m) = \left. \frac{d}{dt} \right|_{t=0} f \left( e^{t\xi} \cdot m \right) . \tag{A.54}$$

For the diagonal $SO(3)$ action, this becomes

$$\xi_M f(\mathbf{p},\mathbf{r}) = \left.\frac{d}{dt}\right|_{t=0} f\left(e^{t\xi}\mathbf{p}, e^{t\xi}\mathbf{r}\right) = \frac{\partial f}{\partial \mathbf{p}}\cdot(\xi\mathbf{p}) + \frac{\partial f}{\partial \mathbf{r}}\cdot(\xi\mathbf{r})$$
$$= \left[(\xi\mathbf{p})\cdot\frac{\partial}{\partial \mathbf{p}} + (\xi\mathbf{r})\cdot\frac{\partial}{\partial \mathbf{r}}\right] f, \tag{A.55}$$

so that

$$\xi_M = (\xi\mathbf{p})\cdot\frac{\partial}{\partial \mathbf{p}} + (\xi\mathbf{r})\cdot\frac{\partial}{\partial \mathbf{r}}\,. \tag{A.56}$$

Plugging this into the relation (2.40),

$$\omega\left(\xi_M,\cdot\right) = -\,\mathrm{d}\Phi^{\xi}(\cdot)\,, \tag{A.57}$$

the left-hand side becomes

$$\omega\left(\xi_M,\cdot\right) = \,\mathrm{d}\mathbf{p} \wedge \mathrm{d}\mathbf{r}\left((\xi\mathbf{p})\cdot\frac{\partial}{\partial \mathbf{p}} + (\xi\mathbf{r})\cdot\frac{\partial}{\partial \mathbf{r}}\,,\cdot\right)$$
$$= (\xi\mathbf{p})\cdot\,\mathrm{d}\mathbf{r} - (\xi\mathbf{r})\cdot\,\mathrm{d}\mathbf{p}\,. \tag{A.58}$$

Checking the right-hand side, we get

$$-\,\mathrm{d}\Phi^{\xi} = -\,\mathrm{d}\left(-\hat{\xi}\cdot(\mathbf{r}\times\mathbf{p})\right) = \hat{\xi}\cdot(\,\mathrm{d}\mathbf{r}\times\mathbf{p} + \mathbf{r}\times\,\mathrm{d}\mathbf{p})$$
$$= \hat{\xi}\cdot(\mathbf{r}\times\,\mathrm{d}\mathbf{p} - \mathbf{p}\times\,\mathrm{d}\mathbf{r})$$
$$= \left(\hat{\xi}\times\mathbf{r}\right)\cdot\,\mathrm{d}\mathbf{p} - \left(\hat{\xi}\times\mathbf{p}\right)\cdot\,\mathrm{d}\mathbf{r}\,. \tag{A.59}$$

Comparing with the previous equation, we need to show

$$\xi\mathbf{p} \overset{\heartsuit}{=} -\left(\hat{\xi}\times\mathbf{p}\right) \qquad \text{and} \qquad \xi\mathbf{r} \overset{\heartsuit}{=} -\left(\hat{\xi}\times\mathbf{r}\right)\,. \tag{A.60}$$

In general, for $\mathbf{A} = \begin{pmatrix} A_x \; A_y \; A_z \end{pmatrix}^T$ one computes

$$\xi\mathbf{A} = \begin{pmatrix} 0 & \xi_z & -\xi_y \\ -\xi_z & 0 & \xi_x \\ \xi_y & -\xi_x & 0 \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = \begin{pmatrix} A_y\xi_z - A_z\xi_y \\ -A_x\xi_z + A_z\xi_x \\ A_x\xi_y - A_y\xi_x \end{pmatrix} = -\hat{\xi} \times \mathbf{A}. \tag{A.61}$$

Hence (A.60) checks out and we have $\omega\left(\xi_M, \cdot\right) = -\,\mathrm{d}\Phi^{\xi}(\cdot)$. This means that expression (A.51) indeed gives a moment map for the diagonal $SO(3)$-action on $\mathbb{R}^6$.

# Appendix B

# The Dimension of $\Phi_1^{-1}\left(\beta_0^*\right)$

The theorem is as follows:[1]

**Theorem 5. Implicit Function Theorem**. *Suppose* $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ *is continuously differentiable in an open set containing* $(a, b)$ *and* $f(a, b) = 0$. *Let* $M$ *be the* $m \times m$ *matrix*

$$\left(D_{n+j} f^i(a, b)\right), \quad 1 \leq i, j \leq m. \tag{B.1}$$

*If* $\det(M) \neq 0$, *there is an open set* $A \subseteq \mathbb{R}^n$ *containing* $a$ *and an open set* $B \subseteq \mathbb{R}^m$ *containing* $b$, *with the following property: for each* $x \in A$ *there is a unique* $g(x) \in B$ *such that* $f(x, g(x)) = 0$. *The function* $g$ *is differentiable.*

So why do we need this? $\Phi_1^{-1}\left(\beta_0^*\right)$ is defined by the relation

$$\Phi_1(\mathbf{r}, \mathbf{p}) = \beta_0^*\,, \tag{B.2}$$

that is

$$\begin{pmatrix} 0 & p_x r_y - r_x p_y & p_x r_z - r_x p_z \\ p_y r_x - r_y p_x & 0 & p_y r_z - r_y p_z \\ p_z r_x - r_z p_x & p_z r_y - r_z p_y & 0 \end{pmatrix} = \begin{pmatrix} 0 & \beta_z & -\beta_y \\ -\beta_z & 0 & \beta_x \\ \beta_y & -\beta_x & 0 \end{pmatrix}. \tag{B.3}$$

[1]For reference, see [53].

This is really only three independent equations:

$$p_y r_z - r_y p_z = \beta_x\,;$$

$$p_z r_x - r_z p_x = \beta_y\,;$$

$$p_x r_y - r_x p_y = \beta_z\,; \tag{B.4}$$

more commonly written as

$$\mathbf{p} \times \mathbf{r} = \hat{\beta}_0\,. \tag{B.5}$$

Then we may define a function $f : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ by

$$(\mathbf{r}, \mathbf{p}) \mapsto \mathbf{p} \times \mathbf{r} - \hat{\beta}_0\,. \tag{B.6}$$

Then $f(\mathbf{r}_0, \mathbf{p}_0) = \mathbf{0}$ for $(\mathbf{r}_0, \mathbf{p}_0) \in \Phi_1^{-1}(\beta_0^*)$. In this case, the matrix of partial derivatives becomes

$$\left(D_{3+j} f^i(\mathbf{r}_0, \mathbf{p}_0)\right) = \left.\begin{pmatrix} \frac{\partial f^1}{\partial p_x} & \frac{\partial f^1}{\partial p_y} & \frac{\partial f^1}{\partial p_z} \\ \frac{\partial f^2}{\partial p_x} & \frac{\partial f^2}{\partial p_y} & \frac{\partial f^2}{\partial p_z} \\ \frac{\partial f^3}{\partial p_x} & \frac{\partial f^3}{\partial p_y} & \frac{\partial f^3}{\partial p_z} \end{pmatrix}\right|_{(\mathbf{r}_0, \mathbf{p}_0)} = \begin{pmatrix} 0 & r_{0z} & -r_{0y} \\ -r_{0z} & 0 & r_{0x} \\ r_{0y} & -r_{0x} & 0 \end{pmatrix}. \tag{B.7}$$

Then we compute

$$\det\left(D_{n+j} f^i(\mathbf{r}_0, \mathbf{p}_0)\right) = 0, \tag{B.8}$$

which we might have expected.

In the Implicit Function Theorem, when $\det\left(D_{n+j} f^i(\mathbf{r}_0, \mathbf{p}_0)\right) \neq 0$, the theorem says there is a neighborhood of $(\mathbf{r}_0, \mathbf{p}_0)$ where the level set of $f$ can be parametrized by $\mathbf{r}$ alone. That is the idea behind the statement $f(x, g(x)) = 0$; the coordinates $x$ alone suffice to describe the level set $A \times B$. The fact that the above equation gives zero determinant suggests that the coordinates $\mathbf{r}$ will not alone be enough to parametrize $\Phi_1^{-1}(\beta_0^*)$.

Instead, rearrange the factors in $(\mathbf{r}, \mathbf{p})$. Let us write

$$(r_x, p_x; r_y, p_y; r_z, p_z) =: (\bar{\mathbf{r}}, \bar{\mathbf{p}}). \tag{B.9}$$

We may define a new function on these coordinates by

$$\bar{f}(\bar{\mathbf{r}}, \bar{\mathbf{p}}) := f(\mathbf{r}, \mathbf{p}); \tag{B.10}$$

*i.e.* $\bar{f}$ has the same functional output as $f$, just with a different arrangement of the input coordinates. Explicitly, we have

$$\bar{f}(r_x, p_x; r_y, p_y; r_z, p_z) = \begin{pmatrix} p_y r_z - r_y p_z - \beta_x \\ p_z r_x - r_z p_x - \beta_y \\ p_x r_y - r_x p_y - \beta_z \end{pmatrix}. \tag{B.11}$$

Then we have

$$\left(D_{3+j}\bar{f}^i(\bar{\mathbf{r}}_0, \bar{\mathbf{p}}_0)\right) = \left.\begin{pmatrix} \frac{\partial \bar{f}^1}{\partial r_x} & \frac{\partial \bar{f}^1}{\partial p_x} & \frac{\partial \bar{f}^1}{\partial r_y} \\ \\ \frac{\partial \bar{f}^2}{\partial r_x} & \frac{\partial \bar{f}^2}{\partial p_x} & \frac{\partial \bar{f}^2}{\partial r_y} \\ \\ \frac{\partial \bar{f}^3}{\partial r_x} & \frac{\partial \bar{f}^3}{\partial p_x} & \frac{\partial \bar{f}^3}{\partial r_y} \end{pmatrix}\right|_{(\bar{\mathbf{r}}_0, \bar{\mathbf{p}}_0)} = \begin{pmatrix} 0 & 0 & -p_{0z} \\ r_{0x} & -r_{0z} & 0 \\ -p_{0y} & r_{0y} & p_{0x} \end{pmatrix}, \tag{B.12}$$

so that

$$\det\left(D_{3+j}\bar{f}^i(\bar{\mathbf{r}}_0, \bar{\mathbf{p}}_0)\right) = -p_{0z}(r_{0x}r_{0y} - p_{0y}r_{0z}) \neq 0 \tag{B.13}$$

for general $(\mathbf{r}, \mathbf{p})$. Hence there is a function $\bar{g}$ such that

$$\bar{f}\left(r_x, p_x; r_y; \bar{g}(r_x, p_x; r_y)\right) = \mathbf{0}, \tag{B.14}$$

saying that the variables $(r_x, p_x; r_y)$ are enough to parametrize the level set of $\bar{f}$ corresponding to the value $\bar{f}(\bar{\mathbf{r}}, \bar{\mathbf{p}}) = \mathbf{0}$.

Notice, in turn, that the level sets of $\bar{f}$ correspond to the level set $\Phi_1(\mathbf{r}, \mathbf{p}) = \beta_0^*$. Hence

$$\Phi_1^{-1}\left(\beta_0^*\right) \quad \text{may be parametrized by} \quad (r_x, p_x; r_y). \tag{B.15}$$

That is,

$$\dim \Phi_1^{-1}\left(\beta_0^*\right) = 3. \tag{B.16}$$

# Appendix C

# A Theorem from Mechanics

For convenience, this section treats a basic theorem from classical mechanics, which states that small changes in the Lagrangian and Hamiltonian are equal in magnitude and opposite in sign.[1] To establish this result, assume a Lagrangian and Hamiltonian are given which depend, not only on the usual coordinates and velocities or momenta, but also on some physical parameter $\lambda$. We then have

$$\begin{aligned} \mathrm{d}L(q,\dot{q},\lambda) &= \sum \frac{\partial L}{\partial q_i}\,\mathrm{d}q_i + \sum \frac{\partial L}{\partial \dot{q}_i}\,\mathrm{d}\dot{q}_i + \frac{\partial L}{\partial \lambda}\,\mathrm{d}\lambda \\ &= \sum \dot{p}_i\,\mathrm{d}q_i + \sum p_i\,\mathrm{d}\dot{q}_i + \frac{\partial L}{\partial \lambda}\,\mathrm{d}\lambda\ . \end{aligned} \tag{C.1}$$

By contrast, since $H = \sum p_i\dot{q}_i - L$, we find

$$\mathrm{d}H(q,p,\lambda) = -\sum \dot{p}_i\,\mathrm{d}q_i + \sum \dot{q}_i\,\mathrm{d}p_i - \frac{\partial L}{\partial \lambda}\,\mathrm{d}\lambda\ . \tag{C.2}$$

This implies that

$$\left(\frac{\partial H}{\partial \lambda}\right)_{p,q} = -\left(\frac{\partial L}{\partial \lambda}\right)_{\dot{q},q}\ , \tag{C.3}$$

where the subscripts denote quantities which are held fixed.

[1]This derivation can be found in [34], §40.

The same result may be phrased in terms of systems and associated perturbations. To see this, suppose $L = L_0 + L_1$, where $L_1$ is a small correction to $L_0$. Let $H_0$ be the Hamiltonian obtained from $L_0$ by Legendre transform, and let $H = H_0 + H_1$ be associated to $L$ in the same fashion. Then the preceding result states that

$$(H_1)_{p,q} = -(L_1)_{\dot{q},q} \ . \tag{C.4}$$

The perturbations, with the appropriate variables held fixed, are equal in magnitude and opposite in sign.

# Appendix D

# Integrals for the Darwin System

There are a few integrals which must be evaluated in simplifying expressions encountered in the Darwin system. These calculations are collected here. Their inclusion in the main body of the text would have distracted too greatly from the main argument at hand. The tedious calculations involved, however, warrant their inclusion, both to save the reader the trouble of calculating them himself, and to allow full-disclosure and a means of checking the results of the line of investigation presented in this thesis.

## D.1 Integrals over $\delta$-Functions

Some integrals involve the use of the Dirac $\delta$-function. These integrals are evaluated below.

### D.1.1 Integral A

Particularly simple is the following:

$$\int_{\mathbf{x}} \frac{\delta^{(3)}\left(\mathbf{x}-\mathbf{q}_a\right)}{\|\mathbf{x}-\mathbf{q}_b\|} = \int \mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3 \frac{\delta\left(x_1-q_{a1}\right)\delta\left(x_2-q_{a2}\right)\delta\left(x_3-q_{a3}\right)}{\sqrt{(x_1-q_{b1})^2+(x_2-q_{b2})^2+(x_3-q_{b3})^2}}$$

$$= \begin{cases} \infty & \text{if } a=b\ ; \\ \frac{1}{\|\mathbf{q}_a-\mathbf{q}_b\|} & \text{if } a\neq b\ . \end{cases} \tag{D.1}$$

### D.1.2 Integral B

The following integral is similar.

$$
\begin{aligned}
I_B &:= \int \mathrm{d}^3\mathbf{x}\delta^{(3)}\left(\mathbf{x}-\mathbf{q}_a\right)\frac{(\dot{\mathbf{q}}_a\cdot\hat{\mathbf{r}}_b)(\dot{\mathbf{q}}_b\cdot\hat{\mathbf{r}}_b)}{\|\mathbf{x}-\mathbf{q}_b\|} \\
&= \int \mathrm{d}x_1\ldots\delta\left(x_1-q_{a1}\right)\ldots\frac{(\dot{q}_{a1}(x_1-q_{a1})+\ldots)(\dot{q}_{b1}(x_1-q_{b1})+\ldots)}{((x_1-q_{b1})^2+\ldots)^{3/2}}\ .
\end{aligned}
\tag{D.2}
$$

If $a=b$, the result is clearly infinite. Otherwise, we find

$$
\begin{aligned}
I_B &= \int \mathrm{d}x_1\delta\left(x_1-q_{b1}\right)\frac{(\dot{q}_{a1}(x_1-q_{b1}))(\dot{q}_{b1}(x_1-q_{b1})}{(x_1-q_{b1})^3} \\
&= \int \mathrm{d}x_1\delta\left(x_1-q_{b1}\right)\frac{\dot{q}_{a1}\dot{q}_{b1}}{x_1-q_{b1}} = \frac{(\dot{\mathbf{q}}_a\cdot\hat{\mathbf{r}}_{ab})(\dot{\mathbf{q}}_b\cdot\hat{\mathbf{r}}_{ab})}{\|\mathbf{q}_a-\mathbf{q}_b\|}\ ,
\end{aligned}
\tag{D.3}
$$

so that

$$
\int \mathrm{d}^3\mathbf{x}\,\delta^{(3)}\left(\mathbf{x}-\mathbf{q}_a\right)\frac{(\dot{\mathbf{q}}_a\cdot\hat{\mathbf{r}}_b)(\dot{\mathbf{q}}_b\cdot\hat{\mathbf{r}}_b)}{\|\mathbf{x}-\mathbf{q}_b\|} = \begin{cases} \infty & \text{if } a=b\ ; \\ \frac{(\dot{\mathbf{q}}_a\cdot\hat{\mathbf{r}}_{ab})(\dot{\mathbf{q}}_b\cdot\hat{\mathbf{r}}_{ab})}{\|\mathbf{q}_a-\mathbf{q}_b\|} & \text{if } a\neq b\ ; \end{cases}
\tag{D.4}
$$

where $\hat{\mathbf{r}}_{ab} := (\mathbf{q}_a-\mathbf{q}_b)/\left\|\mathbf{q}_a-\mathbf{q}_b\right\|$.

## D.2 The Darwin Integral

The next integral to be calculated requires a bit more finesse. Because of its importance in Darwin theory, and because it is the only integral in the theory which requires any argumentation worthy of merit, we may call this '*the* Darwin Integral' without any possibility of confusion. We shall discuss two methods of calculation.

The first method of calculation is actually quite simple once one sees the trick. The trick amounts to converting the integral into an electric potential problem, then using a simple Gauss's Law calculation to obtain the final result. To this end, we first write

$$I_D := \int \frac{\mathrm{d}^3\mathbf{x}}{\|\mathbf{x} - \mathbf{q}_a\| \, \|\mathbf{x} - \mathbf{q}_b\|} = \int \frac{\mathrm{d}^3\mathbf{y}}{\|\mathbf{y} - \mathbf{z}\| \, \|\mathbf{y}\|} =: f(\mathbf{z}), \tag{D.5}$$

where $\mathbf{y} := \mathbf{x} - \mathbf{q}_b$ and $\mathbf{z} := \mathbf{q}_a - \mathbf{q}_b$, and the subscript $D$ stands for 'Darwin'. Then differentiating with respect to $\mathbf{z}$, we obtain

$$\begin{aligned} \nabla^2_{\mathbf{z}} f(\mathbf{z}) &= \int \frac{\mathrm{d}^3\mathbf{y}}{\|\mathbf{y}\|} \nabla^2_{\mathbf{z}} \left( \frac{1}{\|\mathbf{y} - \mathbf{z}\|} \right) = -4\pi \int \frac{\mathrm{d}^3\mathbf{y}}{\|\mathbf{y}\|} \delta^{(3)}(\mathbf{y} - \mathbf{z}) \\ &= \frac{-4\pi}{\|\mathbf{z}\|} =: -4\pi\rho(\mathbf{z}). \end{aligned} \tag{D.6}$$

This is now reminiscent of a potential problem with charge density $\rho(\mathbf{z}) = 1/\|\mathbf{z}\|$. Since this is spherically symmetric, we may work in terms of the radial coordinate $r := \|\mathbf{z}\|$. Enclosing the charge in a sphere of radius $R$, Gauss's Law $\nabla\cdot\mathbf{E} = 4\pi\rho$ gives

$$\begin{aligned} E \int \mathrm{d}a &= \int \mathbf{E} \cdot \, \mathrm{d}a = \int \nabla\cdot\mathbf{E} \, \mathrm{d}^3\mathbf{x} = 4\pi \int \rho r^2 \, \mathrm{d}r \, \mathrm{d}\Omega \\ &= (4\pi)^2 \int_0^R \frac{r^2 \, \mathrm{d}r}{r} = (4\pi)^2 \frac{R^2}{2} \, . \end{aligned} \tag{D.7}$$

Then

$$\mathbf{E}(R) = \frac{1}{2} \frac{(4\pi)^2 R^2}{(4\pi) R^2} \hat{\mathbf{r}} = 2\pi\hat{\mathbf{r}}. \tag{D.8}$$

Computing the potential, and setting the zero point at the origin rather than infinity, we have

$$\phi(R) - \phi(0) = -\int_0^R \mathbf{E} \cdot \, \mathrm{d}r = -2\pi(R - 0), \tag{D.9}$$

or

$$\phi(r) = -2\pi r, \qquad \text{so that} \qquad f(\mathbf{z}) = -2\pi \left\| \mathbf{z} \right\| . \tag{D.10}$$

In terms of our integral, this means

$$\int \frac{\mathrm{d}^3\mathbf{x}}{\|\mathbf{x} - \mathbf{q}_a\| \, \|\mathbf{x} - \mathbf{q}_b\|} = -2\pi \left\| \mathbf{q}_a - \mathbf{q}_b \right\| \qquad \text{for } a \neq b. \tag{D.11}$$

In the case $a = b$, we have

$$\int \frac{\mathrm{d}^3\mathbf{x}}{\|\mathbf{x} - \mathbf{q}_a\| \, \|\mathbf{x} - \mathbf{q}_b\|} = \int \frac{\mathrm{d}^3\mathbf{y}}{\|\mathbf{y}\|^2} = 4\pi \int_0^\infty \frac{r^2 \, \mathrm{d}r}{r^2} = \infty. \tag{D.12}$$

Therefore

$$\int \frac{\mathrm{d}^3\mathbf{x}}{\|\mathbf{x} - \mathbf{q}_a\| \, \|\mathbf{x} - \mathbf{q}_b\|} = \begin{cases} \infty & \text{for } a = b \, ; \\ -2\pi \left\| \mathbf{q}_a - \mathbf{q}_b \right\| & \text{for } a \neq b \, . \end{cases} \tag{D.13}$$

The above result is extremely simple, yet obviously wrong. The left-hand side is the integral of a positive-definite quantity and must therefore be positive. The right-hand side is the negative of a positive-definite quantity, and must therefore be negative. Thus the two can only be equal if they are both zero. A conundrum indeed, since this integral is the crux of Darwin theory, and the calculations in which it is employed certainly require it to have a value other than zero.

Fortunately the above answer is not as wrong as it seems. In a sense it is the only right answer this integral can have if we do not just declare the entire quantity infinite and toss it out altogether. To see in what sense the above calculation produces the right result, we calculate the integral again. This time, however, we do not use tricks (or, if you like, we use tricks of a wholly different sort).

We begin in the same fashion as before, choosing $\mathbf{y} := \mathbf{x} - \mathbf{q}_b$ and $\mathbf{z} := \mathbf{q}_a - \mathbf{q}_b$, so that the integral takes the form

$$I_D = \int \frac{\mathrm{d}^3\mathbf{y}}{\|\mathbf{y}-\mathbf{z}\|\,\|\mathbf{y}\|}. \tag{D.14}$$

We are now free to choose coordinates convenient for performing the integral. Let us therefore place $\mathbf{z}$ along the $y_3$-axis, with $z := \|\mathbf{z}\| > 0$. With this placement of $\mathbf{z}$, if we express $\mathbf{y}$ in spherical polar coordinates as $\mathbf{y} = (r, \theta, \phi)$, then the angle $\theta$ between $\mathbf{y}$ and the $y_3$-axis is also the angle between $\mathbf{y}$ and $\mathbf{z}$. In these coordinates, our the Darwin integral becomes

$$\begin{aligned} I_D &= \int_0^\infty r^2 dr \int_{-1}^1 dx \int_0^{2\pi} d\phi \frac{1}{r\sqrt{r^2+z^2-2rzx}} \\ &= 2\pi z \int_0^\infty r' dr' \int_{-1}^1 dx \frac{1}{\sqrt{r'^2+1-2r'x}}, \end{aligned} \tag{D.15}$$

where $x := \cos(\theta)$ and $r' := r/z > 0$. Our immediate task therefore is to study

$$I(r') := \int_{-1}^1 dx \frac{1}{\sqrt{\beta - \alpha x}}, \qquad \beta := r'^2 + 1, \quad \alpha := 2r'. \tag{D.16}$$

To begin our study of $I(r')$, we first note that $\beta > 0$ and $\alpha > 0$, and moreover that $\beta \geq \alpha$ with equality only when $r' = 1$. This implies that $\beta - \alpha x \geq 0$ for all $x \in [-1, 1]$, and so $I(r')$ is determined up to an overall sign, depending on the choice of positive or negative root $\pm\sqrt{\beta-\alpha x}$. That is, $I(r')$ is either all positive or all negative. In the former case it represents the area under the curve $1/\sqrt{\beta-\alpha x}$, in the latter the area under $-1/\sqrt{\beta-\alpha x}$. In either case, the value of the integrand at the righthand limit of integration $x=1$ is

$$\pm\frac{1}{\sqrt{\beta-\alpha}} \to \pm\infty \qquad \text{as} \qquad r' \to 1. \tag{D.17}$$

For definiteness, we will take the positive root of the integrand.

We now look at the behavior of $I(r')$ as a function of $r'$. It is simple to compute $I(0)=2$, and we also see that, as $r'\to\infty$,

$$I(r') = \int_{-1}^{1} dx \frac{1}{\sqrt{r'^2+1-2r'x}} \sim \int_{-1}^{1} dx \frac{1}{\sqrt{r'^2}} = \frac{2}{r'}, \tag{D.18}$$

so that $I(r')\to 0$ as $2/r'$ when $r'\to\infty$. Using the elementary integral

$$\int \frac{dx}{\sqrt{ax+b}} = \frac{2}{a}\sqrt{ax+b}, \tag{D.19}$$

we may perform the $x$ integral to find

$$I(r') = \int_{-1}^{1} dx \frac{1}{\sqrt{\beta-\alpha x}} = \frac{2}{\alpha}\left[\sqrt{\beta+\alpha}-\sqrt{\beta-\alpha}\right]. \tag{D.20}$$

From this we see that $I(1)=\pm 2$, depending on the root chosen. Since $I(r')$ is either all positive or all negative, and since $I(0)=2>0$, then consistency requires the positive root, hence $I(1)=2$. In general, substituting expressions for $\alpha$ and $\beta$ in terms of $r'$, we have $I(r') = [\pm(r'+1)-(\pm)(r'-1)]/r'$, allowing four possibilities:

$$I(r') = \begin{cases} 2/r' & \text{taking } +,+, \\ 2 & \text{taking } +,-, \\ -2 & \text{taking } -,+, \\ -2/r' & \text{taking } -,-. \end{cases} \tag{D.21}$$

Recalling that we have chosen $I(r')$ to be all positive, we consider only the first two choices as consistent with the calculations done so far. Given that we have shown $I(0) = 2 = I(1)$ and $I(r') \sim 2/r'$ as $r' \to \infty$, then the only choice for $I(r')$ is

$$I(r') = \begin{cases} 2 & \text{for } 0 \leq r' < 1, \\ 2/r' & \text{for } 1 \leq r' < \infty. \end{cases} \tag{D.22}$$

This choice has the virtue of making $I(r')$ a continuous function.

If we now compute $I_D$ as the integral over all $r' > 0$ of $I(r')$, it is clear that we must split the integral into two regions, one for $0 < r' < 1$ and one for $1 \leq r'$. As yet another check on the consistency of our procedure, however, we may step back and compute $I_D$ for all four possibilities of $I(r')$. Over the finite region bounded above by some value $k$, we find

$$\frac{I_D}{2\pi z} = \int_0^k r'dr'I(r') = \int_0^k dr' \begin{cases} 2 \\ 2r' \\ -2r' \\ -2 \end{cases} = \begin{cases} 2k \\ k^2 \\ -k^2 \\ -2k^2 \end{cases} . \tag{D.23}$$

Recall that for large enough $r'$, we know that $I(r') \sim 2/r'$. Thus, for large enough $k$, the the original Darwin integral is

$$
\begin{aligned}
I_D &= 2\pi z \left[\int_0^k r' dr' I(r') + \int_k^\infty r' dr' I(r')\right] \\
&= 2\pi z \left[\begin{cases} 2k \\ k^2 \\ -k^2 \\ -2k^2 \end{cases} + \int_k^\infty dr'\, r' \cdot \frac{2}{r'}\right] \\
&= 2\pi z \left[\begin{cases} 2k \\ k^2 \\ -k^2 \\ -2k^2 \end{cases} + (2 \cdot \infty - 2k)\right]. \qquad \text{(D.24)}
\end{aligned}
$$

From the above expression, we see that a natural way to retain a finite part of the integral is to choose the $k^2$ case above and to set $k = 1$. This amounts to choosing $I(r')$ as in (D.22), so that we find

$$I_D = 2\pi z \left[k^2 + (2 \cdot \infty - 2k)\right]\Big|_{k=1} = 2\pi z[-1 + 2 \cdot \infty]. \qquad \text{(D.25)}$$

Retaining only the finite part of the integral, we find that

$$I_D = -2\pi z, \qquad \text{(D.26)}$$

as we found in (D.13).

# Appendix E

# The Darwin Kernel

We may approach the equivalence of $\mathbf{A}_D$ and $\mathbf{A}_C^{qs}$ in another fashion. Let us recall that the Darwin vector potential solves

$$\nabla^2 \mathbf{A}_D = -\frac{4\pi}{c}\mathbf{j}_t. \tag{E.1}$$

This relates $\mathbf{A}_D$ to the *transverse* current. The current which appears in the Darwin Lagrangian, however, is the *full* current $\mathbf{j}$. In this section we will find that the magnetic field term in the action, written in terms of the Darwin potential $\mathbf{A}_D$

$$-\int \mathrm{d}^3\mathbf{x}\left(B^D\right)^2 = \int \mathrm{d}^3\mathbf{x}\mathbf{A}_D \cdot \nabla^2 \mathbf{A}_D, \tag{E.2}$$

takes the form of the *coupling* term $\int \mathrm{d}^3\mathbf{x}\mathbf{j} \cdot \mathbf{A}$, but where $\mathbf{j}$ is the full (not transverse) current, and $\mathbf{A}$ is the quasistatic potential $\mathbf{A}_C^{qs}$.

To establish this result, we begin by using the above equation for $\mathbf{A}_D$ in terms of the transverse current, and then substitute the explicit form of $\mathbf{j}_t$

to write the following:

$$
\begin{aligned}
&\int \mathrm{d}^3\mathbf{x}\mathbf{A}_D(\mathbf{x},t)\cdot\nabla^2\mathbf{A}_D(\mathbf{x},t)\\
&= -\frac{1}{4\pi c}\int \mathrm{d}^3\mathbf{x}\int \mathrm{d}^3\mathbf{x}''\frac{1}{\|\mathbf{x}-\mathbf{x}''\|}\left[\epsilon_{ijk}\partial_j''\epsilon_{klm}\partial_l''\int \mathrm{d}^3\mathbf{x}'\frac{\mathrm{j}_m(\mathbf{x}',t)}{\|\mathbf{x}''-\mathbf{x}'\|}\right]\\
&\qquad\qquad\cdot\left[\epsilon_{irs}\partial_r\epsilon_{squ}\partial_q\int \mathrm{d}^3\bar{\mathbf{x}}\frac{\mathrm{j}_u(\bar{\mathbf{x}},t)}{\|\mathbf{x}-\bar{\mathbf{x}}\|}\right]\\
&= -\frac{1}{c^2}\int \mathrm{d}^3\mathbf{x}'\mathrm{j}_m(\mathbf{x}',t)\int \mathrm{d}^3\bar{\mathbf{x}}\mathrm{j}_u(\bar{\mathbf{x}},t)K^D_{mu}\left(\mathbf{x}'\mid\bar{\mathbf{x}}\right),
\end{aligned}
\tag{E.3}
$$

where

$$
K^D_{mu}\left(\mathbf{x}'\mid\bar{\mathbf{x}}\right) := \frac{1}{4\pi}\int \mathrm{d}^3\mathbf{x}\int \mathrm{d}^3\mathbf{x}''\frac{\epsilon_{ijk}\epsilon_{klm}\epsilon_{irs}\epsilon_{squ}}{\|\mathbf{x}''-\mathbf{x}'\|\,\|\mathbf{x}-\bar{\mathbf{x}}\|}\partial_r\partial_q\partial_j''\partial_l''\left(\frac{1}{\|\mathbf{x}-\mathbf{x}''\|}\right)
\tag{E.4}
$$

is the Darwin kernel. Again applying the identity $\epsilon_{kij}\epsilon_{klm}=\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl}$, the Darwin kernel may be expressed as

$$
\begin{aligned}
4\pi K^D_{mu}\left(\mathbf{x}'\mid\bar{\mathbf{x}}\right) &= \int \mathrm{d}^3\mathbf{x}\int \mathrm{d}^3\mathbf{x}''\frac{1}{\|\mathbf{x}''-\mathbf{x}'\|\,\|\mathbf{x}-\bar{\mathbf{x}}\|}\left[\partial_u\partial_l\partial_m''\partial_l''-\partial_q\partial_q\partial_m''\partial_u''\right.\\
&\qquad\left.-\partial_u\partial_m\partial_l''\partial_l''+\delta_{mu}\partial_r\partial_r\partial_l''\partial_l''\right]\left(\frac{1}{\|\mathbf{x}-\mathbf{x}''\|}\right).
\end{aligned}
\tag{E.5}
$$

Repeated application of the identity (4.21), along with integration by parts and $\nabla^2(1/\|\mathbf{x}-\mathbf{x}'\|)=-4\pi\delta^{(3)}\left(\mathbf{x}-\mathbf{x}'\right)$, yields

$$
\begin{aligned}
K^D_{mu}\left(\mathbf{x}'\mid\bar{\mathbf{x}}\right) &= \int \mathrm{d}^3\mathbf{x}\int \mathrm{d}^3\mathbf{x}''\frac{-\delta^{(3)}\left(\mathbf{x}-\mathbf{x}''\right)}{\|\mathbf{x}''-\mathbf{x}'\|}\left[-\bar{\partial}_m\bar{\partial}_u+\delta_{mu}\bar{\partial}_r\bar{\partial}_r\right]\frac{1}{\|\mathbf{x}-\bar{\mathbf{x}}\|}\\
&= \frac{4\pi\delta_{mu}}{\|\mathbf{x}'-\bar{\mathbf{x}}\|}-\partial_m'\bar{\partial}_u\int \mathrm{d}^3\mathbf{x}''\frac{1}{\|\mathbf{x}''-\mathbf{x}'\|\,\|\mathbf{x}''-\bar{\mathbf{x}}\|}.
\end{aligned}
\tag{E.6}
$$

We again encounter the Darwin integral, (D.13), so that the above finally becomes

$$K^D_{mu}\left(\mathbf{x}' \mid \bar{\mathbf{x}}\right) = \frac{2\pi}{\|\mathbf{x}' - \bar{\mathbf{x}}\|}\left(\delta_{mu} + \frac{\left(x'_m - \bar{x}_m\right)\left(x'_u - \bar{x}_u\right)}{\|\mathbf{x}' - \bar{\mathbf{x}}\|^2}\right), \tag{E.7}$$

which is a succinct representation of the Darwin kernel matching the form of $\mathbf{A}^{qs}_C$. Using this, we may write

$$\begin{aligned}\int \mathrm{d}^3\mathbf{x}\mathbf{A}_D \cdot \nabla^2 \mathbf{A}_D &= -\frac{2\pi}{c^2}\int \mathrm{d}^3\mathbf{x}' \int \mathrm{d}^3\bar{\mathbf{x}} \frac{\mathbf{j}' \cdot \bar{\mathbf{j}} + (\mathbf{j}' \cdot \hat{\mathbf{r}}')(\bar{\mathbf{j}} \cdot \hat{\mathbf{r}}')}{\|\mathbf{x}' - \bar{\mathbf{x}}\|} \\ &= -\frac{4\pi}{c}\int \mathrm{d}^3\mathbf{x}\mathbf{j} \cdot \mathbf{A}^{qs}_C, \end{aligned} \tag{E.8}$$

the typical form of the Darwin interaction.

# Bibliography

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics.* Addison-Wesley, 1978.

[2] Angel Alastuey and Walter Appel. On the decoupling between classical Coulomb matter and radiation. *Physica*, 276:508–520, 2000.

[3] Walter Appel and Angel Alastuey. Relativistic corrections for a classical one-component plasma with Darwin interactions. *Physica*, 252:238–268, 1998.

[4] V. I. Arnold. *Mathematical Methods of Classical Mechanics.* Springer Verlag, 1989.

[5] Vladimir I. Arnold and Boris A. Khesin. *Topological Methods in Hydrodynamics.* Springer Verlag, 1998.

[6] A.O. Barut and George Craig. The Dynamical Motions of Charged Particles. *Physica*, 197:275–283, 1993.

[7] Timothy H. Boyer. Classical Electromagnetic Interaction of a Point Charge and a Magnetic Moment: Considerations Related to the Aharonov-Bohm Phase Shift. *Foundations of Physics*, 32(1):1–39, 2002.

[8] O. L. Brill and B. Goodman. Causality in Coulomb Gauge. *American Journal of Physics*, 35(9):832–837, 1967.

[9] C.G. Darwin. The Dynamical Motions of Charged Particles. *Reviews of Modern Physics*, 39(233):537–551, 1920.

[10] Ronald C. Davidson, W. Wei-li Lee, Hong Qin, and Edward Startsev. Implications of the electrostatic approximation in the beam frame on the nonlinear Vlasov-Maxwell equations for intense beam propagation. *Plasma Physics*, 9(1):340–344, 2002.

[11] Jayme de Luca. Simple dynamical system with discrete bound states. *Physical Review*, 62(2):2060–2067, 2000.

[12] Luc Dettwiller. Étude perturbative du mouvement de deux particules chargées liées par attraction Coulombienne; effet des termes de Darwin. *European Journal of Physics*, 16:154–163, 1995.

[13] R. L. Dewar. A Lagrangian theory for nonlinear wave packets in a collisionless plasma. *Plasma Physics*, 7(2):267–284, 1972.

[14] Wojciech Frejlak. Generalization of the Darwin Lagrangian. *International Journal of Theoretical Physics*, 27(6):711–716, 1988.

[15] Matthew R. Gibbons and Dennis W. Hewett. Characterization of the Darwin Direct Implicit Particle-in-Cell Method and Resulting Guidelines for Operation. *Journal of Computational Physics*, 130:54–66, 1997.

[16] Robert T. Glassey. *The Cauchy Problem in Kinetic Theory.* Society for Industrial and Applied Mathematics, 1996.

[17] P.P. Goldstein and L.A. Turski. The Vlasov equation for systems with velocity dependent interactions. *Physica*, 89A:481–500, 1977.

[18] P.P. Goldstein and L.A. Turski. On Relativistic Corrections in a Vlasov Plasma. *Physica*, 127A:549–573, 1984.

[19] Robert J. Goldston and Paul H. Rutherford. *Introduction to Plasma Physics.* Institute of Physics Publishing, 1995.

[20] Larry Gonick and Art Huffman. *The Cartoon Guide to Physics.* Harper-Collins, 1992.

[21] Hasan Gümral. Notes on Plasma Physics Geometry, 2001. unpublished.

[22] Peter Havas and John Stachel. Invariances of Approximately Relativistic Lagrangians and the Center-of-Mass Theorem.i. *Physical Review*, 185(5):1636–1647, 1969.

[23] R. D. Hazeltine and S. M. Mahajan. Fluid description of relativistic, magnetized plasma, 2001.

[24] Oliver Heaviside. On the Electromagnetic Effects due to the Motion of Electrification through a Dielectric. *Philosophical Magazine*, 28(167):324–339, 1889.

[25] Oliver Heaviside. On the General Solution of Maxwell's Electromagnetic Equations in a Homogeneous Isotropic Medium, especially in regard to the Derivation of Special Solutions, and the Formulæ for Plane Waves. *Philosophical Magazine*, 28(164):29–50, 1889.

[26] D. W. Hewett. Low-frequency electromagnetic (Darwin) applications in plasma simulation. *Computer Physics Communications*, 84:243–277, 1994.

[27] E. H. Holt and R. E. Haskell. *Foundations of Plasma Dynamics*. Macmillan, 1965.

[28] J. D. Jackson. From Lorenz to Coulomb and other explicit gauge transformations. *American Journal of Physics*, 70(9):917–928, 2002.

[29] J. D. Jackson and L. B. Okun. Historical roots of gauge invariance. *Reviews of Modern Physics*, 73:663–680, 2001.

[30] John David Jackson. *Classical Electrodynamics*. John Wiley & Sons, 1975.

[31] Allan N. Kaufman. Statistical Theory of Magnetic Susceptibility of an Imperfect Gas. *Physical Review*, 136(6A):1614–1625, 1964.

[32] Allan N. Kaufman and Peter S. Rostler. The Darwin Model as a Tool for Electromagnetic Plasma Simulation. *Physics of Fluids*, 14:446–448, 1971.

[33] L. D. Landau and E. M. Lifshitz. *The Classical Theory of Fields.* Elsevier Science, 1975.

[34] L. D. Landau and E. M. Lifshitz. *Mechanics.* Elsevier Science, 1976.

[35] E. Laserra, I.P. Pavlotsky, and M. Strianese. Radius of electron as a consequence of Poincaré group. *Physica*, 219:141–158, 1995.

[36] W. W. Lee and Hong Qin. Alfvén waves in gyrokinetic plasmas. *Plasma Physics*, 10(8):3196–3203, 2003.

[37] W. Wei-li Lee, Hong Qin, and Ronald C. Davidson. Nonlinear perturbative electromagnetic (Darwin) particle simulation of high intensity beams. *Nuclear Instruments and Methods in Physics Research*, 464:465–469, 2001.

[38] A. T. Lin, J. M. Dawson, and H. Okuda. Application of electromagnetic particle simulation to the generation of electromagnetic radiation. *Physics of Fluids*, 17(11):1995–2001, 1975.

[39] Robert G. Littlejohn. Variational principles of guiding centre motion. *Plasma Physics*, 29(1):111–125, 1983.

[40] F. E. Low. A Lagrangian formulation of the Boltzmann-Vlasov equation for plasmas. *Proceedings of the Royal Society of London*, 248:282–287, 1958.

[41] Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*. Springer Verlag, 1994.

[42] Fulvio Melia. *Electrodynamics*. Universtiy of Chicago Press, 2001.

[43] P. J. Morrison. Hamiltonian description of the ideal fluid. *Reviews of Modern Physics*, 70(2):467–521, 1998.

[44] William A. Newcomb. Lagrangian and Hamiltonian Methods in Magnetohydrodynamics. *Nuclear Fusion Supplement*, 2:451–463, 1962.

[45] Clair W. Nielson and H. Ralph Lewis. Particle-Code Models in the Nonradiative Limit. In John Killeen, editor, *Methods in Computational Physics Volume 16, Controlled Fusion*, pages 367–388. Academic Press, 1976.

[46] I. N. Nitikin and J. De Luca. Numerical methods for the three-dimensional two-body problem in the action-at-a-distance electrodynamics. *International Journal of Modern Physics*, 12(5):739–750, 2001.

[47] Yu. N. Orlov. Derivation of the hydrodynamic equations in the framework of the Bogoliubov functional hypothesis. *Theoretical and Mathmatical Physics*, 134(3):439–449, 2003.

[48] Wolfgang K. H. Panofsky and Melba Phillips. *Classical Electricity and Magnetism*. Addison-Wesley, 1962.

[49] Philip L. Pritchett. Particle-in-Cell Simulations of Magnetosphere Electrodynamics. *IEEE Transactions on Plasma Science*, 28(6):1976–1990, 2000.

[50] F. Rohrlich. *Classical Charged Particles: Foundations of Their Theory.* Addison-Wesley, 1965.

[51] Stephanie Frank Singer. *Symmetry in Mechanics: A Gentle, Modern Introduction.* Birkhäuser, 2001.

[52] S. Smale. Topology and Mechanics. *Inventiones Mathematicae*, 10:305–331, 1970.

[53] Michael Spivak. *Calculus on manifolds: a modern approach to classical theorems of advanced calculus.* W. A. Benjamin, 1965.

[54] Walter Thirring. *Classical Mathematical Physics: Dynamical Systems and Field Theories.* Springer Verlag, 1997.

[55] H. Weitzner and W. S. Lawson. Boundary conditions for the Darwin model. *Physics of Fluids*, 1(10):1953–1957, 1989.

[56] John A. Wheeler and Richard P. Feynman. Interaction of the Absorber as the Mechanism of Radiation. *Reviews of Modern Physics*, 17:157–181, 1945.

[57] John A. Wheeler and Richard P. Feynman. Classical Electrodynamics in Terms of Direct Interparticle Action. *Reviews of Modern Physics*, 21(3):425–433, 1949.

[58] Huanchun Ye and P. J. Morrison. Action principles for the Vlasov equation. *Physics of Fluids*, 4(4):771–777, 1992.

# Index

# Vita

Todd Brandon Krause was born in Smithtown, New York on 12 January 1974, the son of Falk H. Krause and Marjorie S. Krause. He received the Bachelor of Arts degree in Physics from Bard College in 1997. While at Bard, he nursed a non-negligible interest in Classical Greek, Latin and Sanskrit. Before leaving college, he spent a year as a performing musician. He started graduate studies in the Physics program at the University of Texas at Austin in August, 1997. During his stay at UT-Austin, his language interests underwent nonlinear growth, resulting in the study of many of the Indo-European languages (including Old Church Slavonic, Old Russian, Hittite, Hindi, Gothic, Old Saxon, Old Norse, Old/Middle High German, Old/Middle English, and Classical Armenian), as well as non-Indo-European (Classical and Yucatec Maya). This resulted in work with Dr. Winfred P. Lehmann at the Linguistics Research Center at UT-Austin contemporaneous with his Physics research.

Permanent address: 5604 Southwest Pkwy.
Apt. 3113
Austin, Texas 78735

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.